Chapter 1 : Introduction

Holography was first conceived almost 50 years ago in 1948 by Dennis Gabor [1] as a method to enhance the resolution of electron microscopes. Today, holographic elements can be found on everyday items such as credit cards, software boxes, baseball cards, book covers and driver's licenses. Holography also has numerous scientific applications in telecommunication and non-destructive testing of stressed materials. Recently, digital holographic data storage has also become a hot topic discussed in popular magazines [2, 3]. Holographic data storage offers two substantial advantages over conventional storage technologies: (1) the ability to store and recall approximately a million bits simultaneously; (2) multiple 'pages' of data can be stored in the same volume and be recalled individually with little cross-talk noise. These improvements translate into a storage device that has potentially a higher data transfer rate and greater storage density than existing devices.

The idea of holographic data storage can be traced as far back as 1963 when Van Heerden first determined the upper limit of the storage density of volume holograms to be V/λ^3 bits/m^3 [4], where V is the volume of the recording material and λ is the wavelength of the light source. A cube with a volume of 1 cm^3 has the potential of storing 1 terabyte

20000526 015

of information, using a laser with a wavelength of 500 nm. However, due to the lack of necessary companion electronic components (spatial light modulators, stable laser sources, large detector arrays, and micro-controllers) at that time, the technology never matured.

Then came the consumer electronics boom of the 1980's. Started by Casio watches, Sony Walkmans and continued into the 1990's with multimedia laptop computers and digital video camcorders. What used to be non-existent or too expensive is now widely available and cheap. Spatial light modulators can be obtained from liquid crystal display panels (LCD) or micro-mirror arrays found in multimedia projection televisions. Stable laser sources in the form of laser diodes are in inexpensive items such as compact disc players and laser pointers. Detector arrays approaching 1,000 by 1,000 pixels are in video camcorders that no tourist would do without. Micro-controllers a thousand times more powerful than the early computers are monitoring everything from car engines to the space shuttles.

More importantly, the need for a large capacity, high data transfer rate storage device is here. The Internet, multimedia video with sound, graphics intensive programs, and video games are overloading computers with data. A minute of full frame MPEG video with sound requires anywhere from a few hundred kilobytes to 60 megabytes of storage space. A 640 by 480 pixels, 24 bit picture is literately worth a thousand kilobytes. The 1996 Olympic Games in Atlanta generated approximately 3 terabytes worth of data from the 271 sporting events and all the participants. End users want this type of data online and they want to access them quickly in order to support the applications running on their computers.

REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

D175

pleting and
acterate for

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REP
		Final 01 Jun 96 to 31 May 99

4. TITLE AND SUBTITLE
Holographic 3D Disks

5. FUNDING NUMBERS
61102F
2305/DX

6. AUTHOR(S)
Professor Psaltis

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
California Institute of Technology
Mail Stop 213-6
Pasadena, CA 91125

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NE
801 N Randolph Street Rm 732
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

F49620-96-1-0196

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
APPROVAL FOR PUBLIC RELEASE, DISTRIBUTION UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The optical correlator and the holographic database can be miniaturized to fit
in samll places such as onboard a vehicle, missile or in the walls of a
building as self contained units.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

The proliferation of consumer electronics and the desire for high capacity and fast data transfer rate storage devices spurred the recent interests in digital holographic data storage. Ironically, some researchers believe that these factors are also narrowing the time frame in which holographic data storage can become successful. It has been predicted based on past performance that hard drive capacity will increase by 60% yearly while the price per megabyte drops by 50% [5]. This allows the hard drive technology to narrow the performance gap between itself and holographic data storage devices while remaining inexpensive to the consumer electronics market. Therefore, in order for holographic data storage devices to compete, it has to be introduced while there is still a substantial performance gap between the two devices. The time window in which this substantial performance gap exists has been predicted to be as short as 2 years (starting from 1996).

However, based on personal observations, I believe this conclusion is erroneous. Holographic data storage devices does not have to compete with hard drives directly. There are other market segments where holographic data storage can prove very useful. For example, CD-ROMs have a user data capacity of only 670 megabytes, way below the average hard drive delivered with computers today (~1 gigabyte). Furthermore, the data transfer rate of CD-ROM players is usually an order of magnitude slower than the average hard drive. By logic, CD-ROM should be an obsolete technology but it is not. Since its introduction in the late 1980's, the compact disc has revolutionized the way information is stored and distributed. For the first time, a randomly accessed, large storage capacity, low cost per megabyte, and removable storage device is available. Unlike hard drives, the compact disc was designed to be an information distribution medium. A compact

disc can be fabricated in less than a minute for less than a dollar, allowing the distribution of large amount of information to the mass market cheaply. The cost of compact disc players are kept low by incorporating only play back features. Some models now allow one time recording on CD-R discs and multiple recordings on CD-RW discs for limited distribution and backups. Nevertheless, the key to the compact disc's success is cost effective distribution of pre-mastered information. First with 74 minutes of high fidelity stereo digital music. Then full motion movies with surround-sound and multiple language tracks with the Digital-Versatile-Disc (DVD).

I believe, there is also a great opportunity for digital holographic data storage devices in the pre-mastered information distribution market. Specifically, the photopolymer based holographic 3 dimensional disk. The following are 4 reasons why I think so:

- **Higher Capacity** - Per volume, holographic data storage offers a much higher storage density than conventional storage devices. Higher storage density translates into higher storage capacity per disk.

- **Faster Data Transfer Rate** - Many bits are recalled in parallel with holography (instead of one bit at a time with compact discs), therefore, the data transfer rate can be considerably higher.

- **Low Cost** - The recording medium, the photopolymer, is composed of inexpensive chemicals that are readily available. Furthermore, mass replication of holographic 3-D disks can be done quickly and cheaply with holographic copying processes where an entire disk is copied in one exposure. The holographic 3-D disk player itself can

be inexpensively made, by borrowing existing components from compact disc players.

- **Removablility** - Similar to the compact disc, the photopolymer based holographic 3-D disk is designed to be an information distribution medium. Therefore, the holographic 3-D disk can be moved from player to player to transfer large quantity of information.

With better performance, low cost per disk, and removability, it is my hope that photopolymer based holographic 3-D disks can replace the compact disc format as the next information distribution medium standard.

1.1 Outline of Thesis

This thesis presents an experimentalist's approach to data storage and information processing using volume holography. Therefore, some background information and theoretical derivations have been omitted in favor of more concise explanations of the experimental setups and results. Most of the theoretical analysis in this thesis have been reduced to very simple forms using diagrams. To get a more complete picture of what holographic data storage is all about, I strongly recommend several other Caltech Ph.D. thesises [6-8].

Chapter 2 discusses in detail the characteristics of the DuPont photopolymer. In order to use the DuPont photopolymer as a storage medium, we fully characterized its recording behavior by using a new multiplexing method called peristrophic multiplexing. For relatively thin materials such as the photopolymer, peristrophic multiplexing gives a

non-Bragg effect method of storing more holograms. From the recording behavior, we were able to derive an exposure schedule to multiplex equal strength holograms in the same volume, using the entire dynamic range of the photopolymer. Peristrophic multiplexing and the exposure schedule were experimentally demonstrated in an experiment where 1,000 equal strength holograms were multiplexed in a 100 microns thick photopolymer.

In Chapter 3, results from several different high density holographic data storage experiments are presented to demonstrate that information can be stored more densely using holography. We start with a preliminary experiment to show that the DuPont photopolymer could indeed be used to record high density holograms with sufficient SNR and end with an experiment where a surface density of 100 bits/μm^2 was achieved. For one of the high density experiments, we constructed a functional holographic 3-D disk record/read station. A new multiplexing method called 'shift' multiplexing was used to write overlapping holograms in a spiral track on the disk, just like the compact disc format. The capacity of a 120 mm disk with a surface density of 100 bits/μm^2 would be around 100 Gbytes, roughly 100 times more than the current compact disc. We end Chapter 3 with some suggestions for future experiments in the area of holographic 3-D disks.

Holography can also be used to process a vast amount of information very quickly through the parallel nature of optics. In Chapter 4, we show optical correlator systems that use a holographic database to navigate an autonomous vehicle, to recognize/track targets, and to identify users through sequential fingerprint inputs. These complicated

tasks were all performed in real-time, demonstrating the capability of photopolymer based optical correlator systems.

References

[1] D. Gabor, "A new microscope principle," *Nature*, 161:777, 1948.

[2] D. Psaltis and F. Mok. "Holographic memories," *Scientific American*, 273(5):70-76, Nov 1995.

[3] J. F. Heanue, M. C. Bashaw, and L. Hesselink. "Volume holographic storage and retrieval of digital data," *Science*, v265(5173):749-752, Aug 1994.

[4] P. J. van Heerden. "Theory of optical information storage in solids," *Applied Optics*, 2(4):393-401, 1963.

[5] "Market Watch," *PC Today*, http://www.pc-today.com.

[6] H.-Y. S. Li. *Photorefractive 3-D disks for optical data storage and artificial neural networks*. Ph.D. thesis, California Institute of Technology, 1994.

[7] K. Curtis. *3-D photopolymer disks for correlation and data storage, and cross-talk in volume holographic memories*. Ph.D. thesis, California Institute of Technology, 1994.

[8] G. W. Burr. *Volume holographic storage using the 90° geometry*. Ph.D. thesis, California Institute of Technology, 1996.

Chapter 2 : Holographic Data Storage Using DuPont's Photopolymer

2.1 Introduction

There are many materials in which holograms could be stored. Early holographers used sliver halide emulsion plates to record both transmission and reflection planar holograms. Some of those films were specially developed for holography and they are still in production. They have very good resolution (5,000 lines/mm) and a fast speed (20 $\mu J/cm^2$) [1]. Like normal photographic films, the exposed film has to be developed with chemicals in order to reveal the stored hologram. Because the hologram is very thin, it can be viewed from a wide range of angles. Hence only one hologram can be stored per location or else they would become indistinguishable.

Another recording material suitable for holography was discovered by accident in the late 1960's [2]. Researchers working on second-harmonic generation with ferroelectric inorganic crystals found that the index of refraction of the crystal changed as it absorbed light. This was later termed the photorefractive effect and one of the crystals exhibiting this effect is lithium niobate ($LiNbO_3$). Since the index of refraction of $LiNbO_3$ could be changed locally by shining a spatially varying light of the proper

wavelength, it wasn't long before people started storing holograms in them [3]. Holograms written in photorefractive crystals are self-processing (no chemical processing is required to reveal the holograms) and since the material is often thick, the viewing angle of the hologram is very limited. This means that a number of holograms could be stored in the same volume and recalled independently by viewing it from different angles. Being able to store many holograms in the same volume significantly increases the storage capacity of the recording material. In a recent experiment, a system capable of storing 160,000 hologram in a block of $LiNbO_3$ was demonstrated [4].

$LiNbO_3$ possess some very nice properties for holographic data storage. It can be grown to thickness varying from tens of micron to a few cm. It can be polished to very good optical quality by standard polishing techniques. Since its Curie temperature is very high (1100 $^{\circ}$C), it can be anti-reflection (AR) coated by vacuum deposition methods which usually yield the best quality. One property that could be both good and bad is that subsequent illumination of the crystal causes the strength of the previously recorded holograms to decay. This can be good because the holograms can be erased and the crystal reused. This can also be bad because the process of reading a hologram also erases slowly, all the stored holograms. On the average, it is a bad property since it is often desirable to read the stored holograms many cycles without decaying their strength. Furthermore, holograms usually do not erase completely with just illumination (baking in an oven is often required to renew the crystal). There are several methods with which to 'fix' the holograms or to delay the decay of the stored holograms [5-8]. These methods usually require some post processing steps, multiple laser sources, or continuous hologram refreshing.

Some other undesirable properties of LiNbO$_3$ are cost (it has to be grown and polished), sensitivity to shock (g-forces and environmental), anisotropy (index of refraction is different along different axis), and slow speed (low sensitivity). Since the focus of this research is to construct a holographic 3-D disk system that is low cost and capable of being read over and over again. LiNbO$_3$ is not well suited for this application.

A better candidate for holographic 3-D disks is the photopolymer [9]. Exposure to light causes an irreversible chemical reaction in which the density of the illuminated area changes. Unlike photographic film, the photopolymer is self-processing and the hologram is available shortly after exposure. The photopolymer is isotropic, inexpensive, and easy to apply to large surfaces, allowing for the fabrication of cheap distributive disk-type devices. On the other hand, the photoplymers are usually fairly thin (10 - 100 μm), limiting the number of holograms that could be stored in the same volume when conventional holographic multiplexing methods are used. The induced change in index of refraction is caused by polymerization of monomers. This causes the photopolymer to shrink slightly during recording, which can affect the quality of the recalled data. Polaroid recently developed a new type of photopolymer that exhibits minimal shrinkage when exposed to light [10]. We will come back and discuss the properties of a particular photopolymer, the DuPont photopolymer, in detail in Chapter 2.2

More recently, there has been some development in hybrid materials called photorefractive polymers [11]. Instead of being grown like a crystal, the photorefractive components are embedded in a polymer host. This eliminates the restriction of a crystal lattice and allows incompatible components to be mixed in the same volume. The characteristics of the photorefractive polymer could be fine-tuned by using different

materials and concentrations. So far, holograms with diffraction efficiencies approaching 100% and a net two-wave coupling gain of 200 cm^{-1} have been reported using this type of materials [12].

Since photorefractive polymer consists of both photorefractives and polymers, it shares some of their weaknesses. Due to the photorefractive effect, the holograms decays upon readout. Furthermore, since the mobile charges in photorefractive polymers are holes, a large electric field has to be applied during recording to form the necessary space charge distribution. Finally, because it is polymer based, the optical quality degrades with increasing thickness. Therefore, only thin samples of the photorefractive polymer have been examined so far.

2.2 DuPont Photopolymer Basics

DuPont has over 20 years of experience in manufacturing photopolymers for holography [13, 14] and currently produces a complete line of OMNIDEX^{TM} films covering the wavelength range from 450 to 660 nm. The OMNIDEX^{TM} films are completely self-processing and the holographic images form in real time. Since it is a commercial product, the final hologram has been tested at high temperatures with humidity and solar UV exposure without noticeable degradation in quality. Applications of these films include: holographic data storage, head-up-display in cars and planes, fiber optic demultiplexers, helmet mounted displays, holoarts, micro-optics, high resolution lithography, Fourier-plane matched filters for pattern recognition, etc.

The recipe for making DuPont's photopolymer is fairly simple and the ingredients include: solvents, polymeric binders, acrylic monomers, plasticizers, photosensitizing dyes, initiating systems, and chain transfer agents. These ingredients are mixed and then dissolved in the solvent. The polymeric binders serve to hold the goo together after it is cast onto 200D Mylar® polyester film using conventional web coaters and the solvents allowed to evaporate. A thin protective cover sheet (usually Mylar film) is then applied to the tacky coating for compete coverage. DuPont's photopolymer is shipped to the customer in rolled drums, page sized sheets, or in some instances, in the original solution form. The solution form is useful if the customer wants to cast or spin coat on his or her own substrates.

Figure 2.1 shows how illumination causes the index of the photopolymer to change. Two laser beams intersect in the photopolymer and setup a sinusoidal interference pattern. The photosensitizing dye absorbs the incident photons and triggers the photoinitiator system to begin photopolymerization of the acrylic monomers (several different photosensitizing dyes could be mixed in one photopolymer to achieve wide color sensitivity). The consumption of monomers in the bright region sets up a monomer concentration gradient, causing net diffusion of monomers from the dark into the bright regions. Polymerization and diffusion of the monomers continue until either the illumination source is removed, or all the monomers are used up. The local index variation in the photopolymer results from the increased density in the bright regions.

Three to four steps are usually involved in using the DuPont photopolymer for holography: (1) laminate or spin coat the photopolymer to a supporting substrate, (2) exposure to laser illumination in the proper wavelength range, (3) a uniform UV exposure

to cure the material of excess monomers, (4) heating the cured photopolymer to increase index modulation. Step 4 is recommended for some DuPont photopolymer but not required. The heat softens the film and allows further monomer diffusion. This can sometimes increase the index modulation by a factor of 2 to 3 times [15].
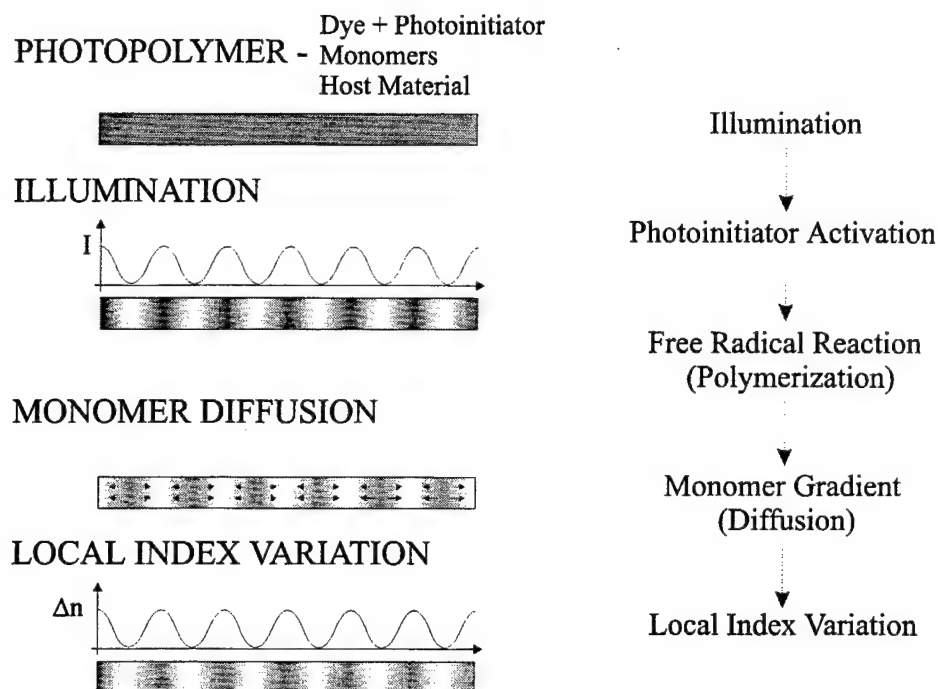
PHOTOPOLYMER - Dye + Photoinitiator
Monomers
Host Material

ILLUMINATION

I

MONOMER DIFFUSION

LOCAL INDEX VARIATION

Δn

Illumination

↓

Photoinitiator Activation

↓

Free Radical Reaction
(Polymerization)

↓

Monomer Gradient
(Diffusion)

↓

Local Index Variation

Figure 2.1 : Photopolymer basics. A sinusoidal illumination causes local index variation in the photopolymer.

One problem with the OMNIDEX^{TM} series photopolymer is their limited thickness of 10 to 25 μm. Unlike photorefractive crystals, the DuPont photopolymers cannot be polished because it is soft and tacky (before complete polymerization). Therefore, in order to maintain good optical quality, the film is made very thin. DuPont also produces an experimental-grade photopolymer in 38 and 100 micron thick versions (HRF-150-38 and HRF-150-100). They are sensitive to the blue-green wavelengths and behave similarly to the OMNIDEX^{TM} films. However, even at 100 micron thick, only a few

holograms can be multiplexed in the same volume and be separable from one another using conventional multiplexing methods. Therefore, before these thin photopolymers can be useful for holographic data storage, we need to figure out a way to squeeze more holograms in the same volume.

2.3 Peristrophic Multiplexing

2.3.1 Introduction

The number of holograms that can be multiplexed in a given holographic system is primarily a function of two parameters - the system's design ($f/\#$ of the lenses, complexity of the data, hologram multiplexing method(s), the thickness of the recording material, laser power, sensitivity of the detectors) and the recording material's dynamic range (diffraction efficiency per hologram). Since DuPont's photopolymers have a relatively high dynamic range, the number of holograms that could be stored is usually limited by the system's design.

The $f/\#$ of the lenses, laser power, detector sensitivity, and the complexity of the data to be stored (number of pixels per page and the pixel size) are usually fixed to suit the specifications of a particular design. The thickness of DuPont's photopolymer is limited to 38 and 100 micron thick in the easy to use sheet form. Therefore, the only thing we could improve is the multiplexing method. One sure way to increase the number of holograms stored is to multiplex many holograms in one location and then repeat the process at many spatial locations (spatial multiplexing). However, in order to maximize the total number of holograms stored, we would like to multiplex as many

holograms in one location as possible before moving to spatial multiplexing. Using angle multiplexing, up to 10 holograms have been stored in a 38 micron thick photoplymer at the second null [16]. Angle multiplexing was achieved by rotating the recording material around the vertical axis while the plane of interaction is in the horizontal direction. The holograms were separated by 2°, requiring a total of 18° in rotation to store the 10 holograms. Compared with the number of holograms that have been stored in thick samples of $LiNbO_3$ [17-19], 10 holograms is a very small number.

The idea for peristrophic multiplexing [20] came to me by accident. I was reading a paper about holographic 3-D disks written by a fellow group member [21] when an interesting question popped into my head. What if a hologram was stored at the exact center of the disk and the disk was rotated during reconstruction. What would happen to the reconstruction? I posed this question to Kevin Curtis and instead of laughing it off, he gave it some serious thought. Since we were used to thinking of the κ-sphere in only two dimensions (in the plane of interaction), we couldn't come to an agreeable conclusion to the question posed. Instead, we tried it in the lab and discovered some very interesting results.

2.3.2 Peristrophic Theory

Figure 2.2 shows a Fourier plane peristrophic multiplexing setup. The data presented on the spatial-light-modulator (SLM) is imaged to the detector array by a 4-f system. The photopolymer is mounted on a rotational stage and placed near the Fourier plane of the 4-f system. The signal and the reference beams make an angle of θ_s and θ_r with respect to the surface normal of the photopolymer. A hologram is recorded in the

photopolymer when both the signal and reference beams illuminate the area at the center of the rotation. To reconstruct the stored hologram, only the reference beam is used to illuminate the photopolymer.
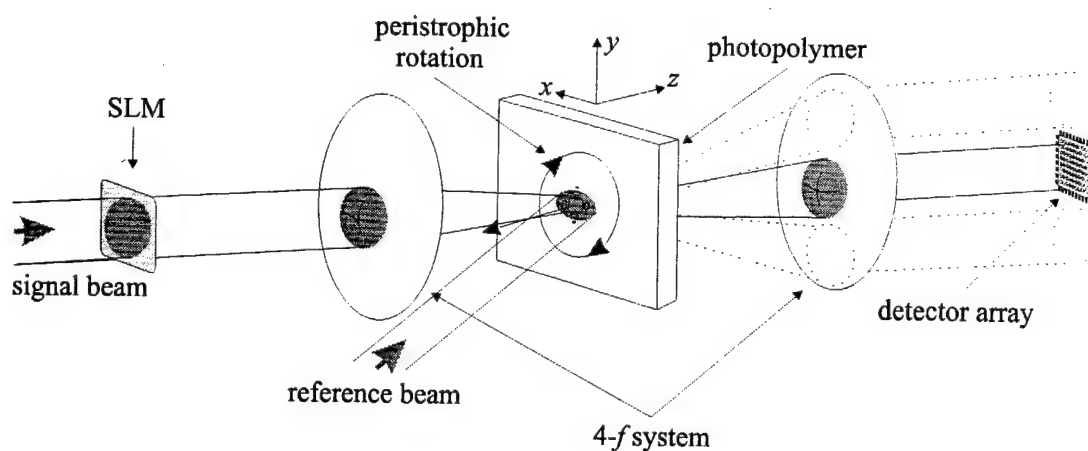

Figure 2.2 : Fourier plane peristrophic multiplexing setup.

When the photopolymer is rotated during reconstruction, the reconstructed image moves in the trajectory of a circle, with the reference beam at the center of the circle. Therefore, by rotating the recording material, we can shift the reconstruction from the previously stored hologram off of the detector array, allowing another hologram to be multiplexed in the same area with very little cross-talk. This is great because even though the reconstruction intensity of the previous hologram might still be very strong, the detector array does not see it. It is as if the previously stored hologram is not even there. Peristrophic multiplexing is simply the rotation of the recording material between holograms, with the axis of rotation being perpendicular to the material's surface. With peristrophic multiplexing, we are no longer relying on the Bragg condition to diminish the intensity of the undesired reconstruction before another hologram could be stored.

To better explain the trajectory of the reconstructed image as the recording material is rotated, let us look at the much simpler setup shown in Figure 2.3. In Figure 2.3 (a), two planewaves intersect to record a hologram at the center of peristrophic rotation. When the recording material is rotated, the holographic grating also rotates at the same rate as the material. During reconstruction, the reference beam illuminates the rotated grating, causing the reconstruction to deflected at a different angle. Since the grating is at the center of the rotation, the trajectory of the reconstructed hologram forms a circle with the reference beam at the center.

The same conclusion can be reached by looking at the κ-sphere diagram in Figure 2.3 (b). The reference and signal beam vectors interfere and record a grating vector. The rotation of the grating is depicted in the κ-sphere diagram as a rotation of the grating vector around the tip of the reference beam vector. The tip of the grating vector traces out a circle in the x-y plane with the reference beam vector at the center.

(a) (b)
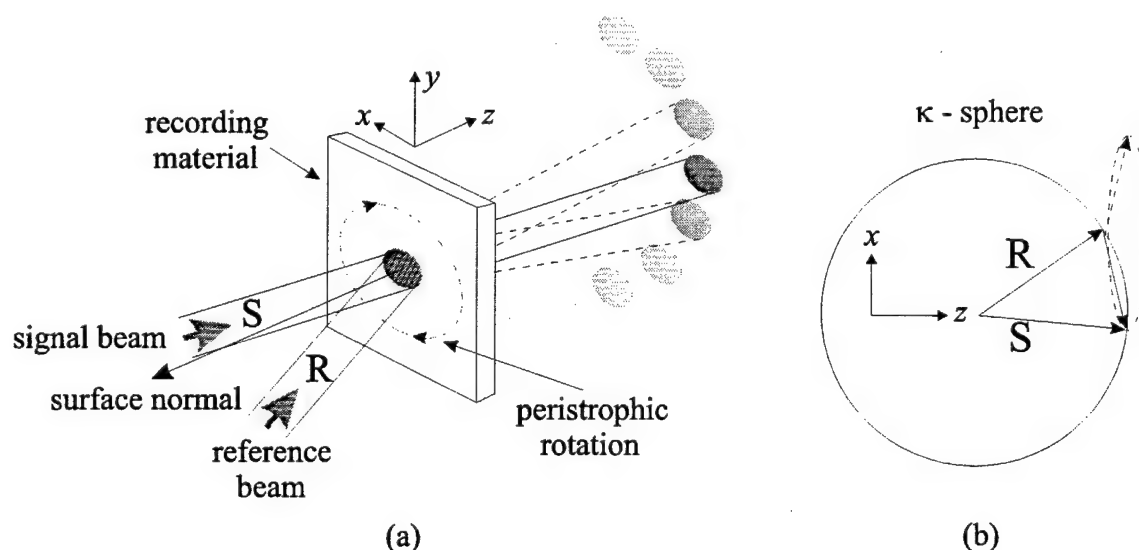
Figure 2.3 : A simpler diagram of peristrophic multiplexing. (a) The recording geometry. (b) The κ-sphere representation.

Figure 2.3 (b) also shows that as the grating vector is rotated around the tip of the reference beam vector, the tip of the grating vector actually leaves the κ-sphere. This means that the reconstructed hologram not only shifts as the material is rotated, its intensity also decreases due to the Bragg effect. For infinitely thin materials, the Bragg effect is not noticeable and the reconstructed hologram remains strong throughout its trajectory. For very thick materials, the Bragg effect is more dominant and the reconstructed hologram might disappear completely before tracing out much of its trajectory. An interesting effect occurs when the material is very thick and the signal bandwidth is high. A combination of Bragg and shift is required to get rid of the undesired reconstruction. To see the relationships between them, the amount of rotation required to shift the reconstructed image off of the detector array and the amount of rotation required to Bragg mis-match the reconstructed image were derived.

The Bragg selectivity can be computed by assuming Born and paraxial approximations and then integrating over the volume of the hologram. With the plane of interaction in x-z plane (Figure 2.3 (b)), the signal and reference beams can be expressed as :

$$\text{reference planewave} \quad R = e^{-ik(\sin \theta_r x + \cos \theta_r z)} \tag{2.1}$$

$$\text{signal planewave} \quad S = e^{-ik(-\sin \theta_s x + \cos \theta_s z)} \tag{2.2}$$

For small peristrophic rotations, the output coordinate, x' and y', is related to the hologram coordinate, x and y, by $x' \approx x - y d\psi$ and $y' \approx y + x d\psi$, where $d\psi$ denotes the peristrophic rotation angle. By assuming that the transverse (x-y plane) dimensions of the

film are much larger than the bandwidth of the signal beam, the Bragg selectivity for the transmission geometry can be shown to be:

$$d\psi = \sqrt{\frac{2\lambda}{L}\frac{\cos\theta_s}{\sin\theta_r(\sin\theta_r + \sin\theta_s)}} \qquad (2.3)$$

where $d\psi$ denotes the peristrophic rotation required to reach the first null, λ is the wavelength of the laser, L is the thickness of the recording material, θ_s and θ_r are the signal and reference beam angles with respect to the material's surface normal (for Bragg selectivity, all terms are referred to inside of the recording material). For example, using $\lambda = 500$ nm, $L = 38$ μm, $\theta_s = \theta_r = 30°$, and letting the index of refraction of the recording material $= 1$, a peristrophic rotation of $d\psi = 12°$ is required to reach the first null.

How much would the reconstruction shift during the 12° in rotation? Perhaps the reconstruction is completely off of the detector array with less $d\psi$ than required for Bragg mis-match. Figure 2.4 shows the effect of peristrophic rotation on the reconstructed image.
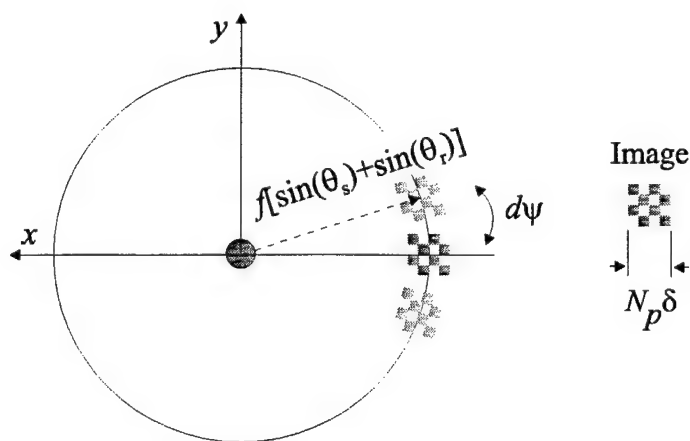
Figure 2.4 : Diagram showing the effect of peristrophic rotation on shifting the reconstruction (Fourier plane recording).

An image with N_p number of pixels in one dimension and a pixel width of δ is recorded in the Fourier plane with the setup shown in Figure 2.2. The total height of the reconstructed hologram is $N_p\delta$ and is readout by a detector array with the same aperture size. The period of the recorded grating along the x axis is equal to :

$$\Lambda = \frac{\lambda}{\sin \theta_r + \sin \theta_s} \qquad (2.4)$$

and the reconstructed hologram is centered at the detector plane at $x = \lambda f / \Lambda$, where f is the focal length of the imaging lens and Λ is the period of the recorded grating. Therefore, the radius of the circle (with the reference beam at the center) on which the reconstructed hologram lies as the material is peristrophically rotated is equal to $f(\sin\theta_r + \sin\theta_s)$. In order to shift the reconstructed hologram completely off of the detector aperture, the following minimum rotation is required :

$$d\psi = \frac{\dfrac{N_p\delta}{f}}{\sin \theta_r + \sin \theta_s} \qquad (2.5)$$

Notice Eq. 2.5 is independent of L, the recording material's thickness. If we plug in $N_p\delta = 1$ cm, $f = 10$ cm, $\theta_s = \theta_r = 30°$, then we get a $d\psi$ of only 5.7°. Therefore for these parameters, a peristrophic rotation of only 5.7° is required to get rid of the undesired reconstruction. In Equation 2.5, $N_p\delta/f$ is approximately the inverse $f/\#$ (the bandwidth) of the signal beam. The higher the bandwidth (lower $f/\#$) the more peristrophic rotation is required to shift the image off of the detector. For example, instead of $N_p\delta = 1$ cm, let $N_p\delta = 5$ cm and the other parameters remain the same, then the amount of peristrophic rotation required to shift the reconstruction off of the detector array becomes 28.6°.

Compared with Bragg peristrophic rotation of only 12° before, one might think that when the recording material is peristrophically rotated by 12°, the reconstruction would be Bragg mis-matched and disappear. Therefore, there would be no need to rotate the recording material the full 28.6° to get rid of the reconstruction. Sadly this is not the case. The peristrophic Bragg selectivity for the transmission geometry (Eq. 2.3) was derived for the case where the signal and reference beams are planewaves. With $N_p\delta = 5$ cm and $f = 10$ cm, this corresponds approximately to a $f/2$ lens with a bandwidth of ~28 degrees. In order to Bragg mis-match the entire signal beam spectrum, the actual peristrophic rotation required is much higher.

A simple way to look at how much rotation is required for high bandwidth holograms is shown in Figure 2.5. In Figure 2.5 (a), the planewave reference beam is represented as a single vector while the high bandwidth signal beam is represented as a spectrum of vectors (also called the signal cone). There exists a grating vector from the tip of the reference beam vector to every signal beam vector in the signal cone. One such grating vector is drawn in Figure 2.5 for illustration. In Figure 2.5 (b), we are looking at the same κ-sphere except in the x-y plane. The signal cone is now represented as a semi-circular patch of area on the κ-sphere and each point in the patch has a grating vector associated with it.

By rotating the recording material peristrophically, all the grating vectors also rotate. Imagine Figure 2.5 (b) as a 3D graph, like half of a sphere sticking out from the page. On this sphere, the circular patch of grating vectors moves as the recording material is rotated. Since the patch resides on a sphere, when moved as shown in Figure 2.5 (b), the top half of the patch lifts off of the κ-sphere while the bottom half dips below

the κ-sphere. There exists a curved line on which the patch intersects the κ-sphere. This line represents a strip of hologram that is exactly Bragg matched and is deflected toward the detector array. We need to continue the peristrophic rotation until the patch is lifted completely off of the κ-sphere. This turns out to be the same amount of peristrophic rotation required to shift the entire spectrum off of the detector array. Therefore with peristrophic multiplexing, regardless of the thickness of the recording material, the minimum rotation required to get rid of the undesired reconstruction is always dictated by Eq. 2.5 (when recorded in the Fourier plane and in the transmission geometry).

(a)   (b)
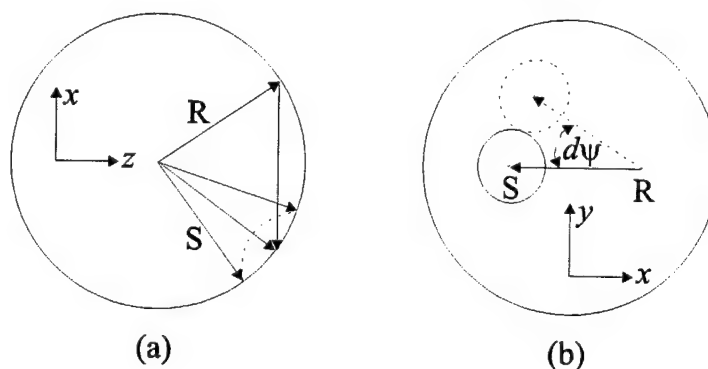
Figure 2.5 : κ-sphere representation of a high bandwidth hologram. (a) Viewed from the x-z plane. (b) Viewed from the x-y plane.

Peristrophic multiplexing can also be applied to image plane holograms. Figure 2.6 shows an image plane peristrophic multiplexing setup. Due to the imaging system in the image plane setup, we cannot shift the undesired reconstructions off of the detector array. However, it is possible to block the shifted reconstructions at the Fourier plane of the second 4-f imaging system. Peristrophic rotation shifts the reconstructed images in the Fourier plane. By placing a spatial filter with the appropriate aperture in the Fourier plane, we can block the undesired holograms.
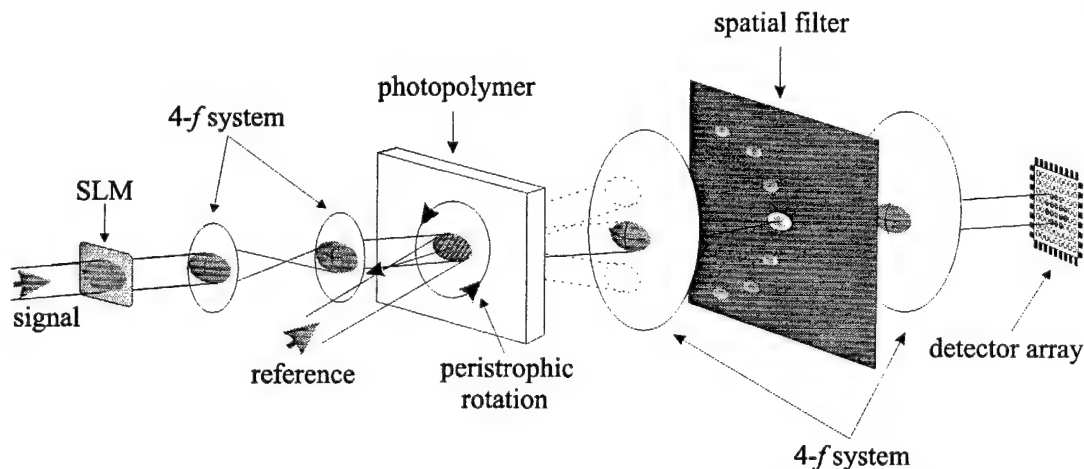
Figure 2.6 : Image plane peristrophic multiplexing setup.

Figure 2.7 shows the graphical representation of the filtering process at the Fourier plane. An image with a pixel width of δ is recorded in the image plane with the setup shown in Figure 2.6. At the Fourier plane, the width of the zero[th] order of the reconstructed image is $2\lambda f/\delta$. By making the aperture of the spatial filter at the Fourier plane equal to the width of the Fourier transform of the reconstructed hologram, an unrotated hologram can pass through and be read out by the detector. Rotating the recording material peristrophically causes the reconstruction to shift and become blocked by the spatial filter, allowing another hologram to be stored. Once again the radius of the circle on which the Fourier transform lies is equal to $f(\sin\theta_r + \sin\theta_s)$. In order to shift the reconstructed hologram completely off of the spatial filter aperture, the following minimum rotation is required :

$$d\psi = \frac{\dfrac{2\lambda}{\delta}}{\sin\theta_r + \sin\theta_s} \tag{2.6}$$

where δ is the pixel size of the image. Notice Eq. 2.6 is also independent of L, the recording material's thickness. If we plug in $\delta = 50$ μm, $\lambda = 500$ nm, $\theta_s = \theta_r = 30°$, then we get a $d\psi$ of only $1.2°$
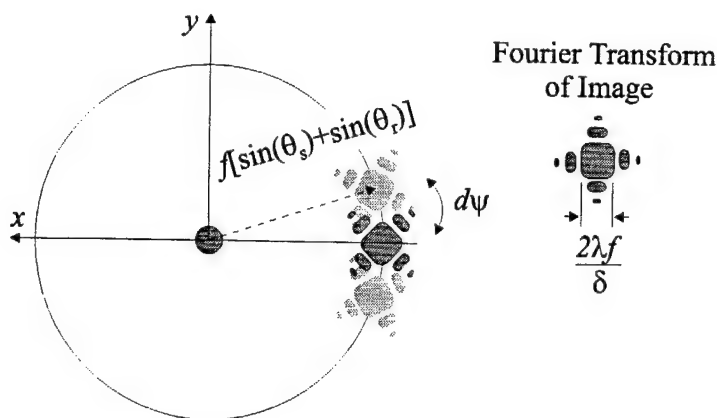

Figure 2.7 : Diagram showing the effect of peristrophic rotation on shifting the reconstruction (image plane recording).

Once again the amount of peristrophic rotation required to shift the undesired reconstructions onto the spatial filter depends on the bandwidth of the signal beam. This time the bandwidth is encoded in δ, the size of the pixels in the stored image. For example, if the pixel size is 5 microns instead of 50 microns, then the size of the Fourier transform would be ten times larger and $d\psi$ would also increase by ten folds to $12°$. Furthermore, we can relate the pixel size to the $f/\#$ of the imaging lenses. It turns out that the minimum resolvable pixel size is approximately equal to $2\lambda f/\#$. By substituting δ in Eq. 2.6 with $2\lambda f/\#$ you will see that Eq. 2.6 is exactly the same as Eq. 2.5. This means if the pixel size of the input image is picked to be at the resolvable limit of the lenses, then the amount of peristrophic rotation required to get rid of the undesired hologram in the image plane setup is the same as the Fourier plane setup.

The above discussion applies to transmission geometry holograms only. The same calculations can be done for reflection geometry holograms as well. It turns out there is only a slight sign change in the grating period. Table 2.1 gives a summary.

	Transmission Geometry	Reflection Geometry
Bragg Selectivity	$d\psi = \sqrt{\dfrac{2\lambda}{L}\dfrac{\cos\theta_s}{\sin\theta_r(\sin\theta_r + \sin\theta_s)}}$	$d\psi = \sqrt{\dfrac{2\lambda}{L}\dfrac{\cos\theta_s}{\sin\theta_r(\sin\theta_r - \sin\theta_s)}}$
Image Plane Recording	$d\psi = \dfrac{\dfrac{2\lambda}{\delta}}{\sin\theta_r + \sin\theta_s}$	$d\psi = \dfrac{\dfrac{2\lambda}{\delta}}{\sin\theta_r - \sin\theta_s}$
Fourier Plane Recording	$d\psi = \dfrac{\dfrac{N_p\delta}{f}}{\sin\theta_r + \sin\theta_s}$	$d\psi = \dfrac{\dfrac{N_p\delta}{f}}{\sin\theta_r - \sin\theta_s}$

Table 2.1 : A summary of peristrophic selectivity for transmission and reflection geometry hologram.

One question remains, can we continuously multiplex peristrophic holograms from 0° to 360° in rotation? The answer is: it depends on the setup. For example, let us look at Figure 2.8 (a). Two planewaves are symmetric with respect to the surface normal of the recording material record a grating. The grating has a symbolic arrow pointing from the tip of the reference beam vector to the tip of the signal beam vector. This just means that when the reference beam is introduced during reconstruction, the tip of the grating vector shows where the signal beam will exit. As the recording material is peristrophically rotated, the tip of the grating vector lifts off of the κ-sphere, causing the reconstruction to become Bragg mis-matched. If the recording material is thick, the reconstruction will disappear very quickly. When the recording material is

peristrophically rotated by 180°, the grating vector is sticking way out of the κ-sphere as shown in Figure 2.8 (b). However, since the arrow on the grating vector is only symbolic (there is no difference between the tip and the tail of the grating vector), we can also put the tip of the grating vector at the tip of the reference beam (Figure 2.8 (c)). For this case, the tail of the grating vector rest on the κ-sphere, and therefore it is Bragg matched. Hence for symmetric signal and reference beams, holograms can only be peristrophically multiplexed from 0° up to (but not including) 180°. At 180° rotation, the reconstruction from the 0° degree hologram will be strong and deflected directly toward the detector array.

For asymmetric planewave recording setup, depending on the thickness of the recording material, up to the full range of 360° can be used for peristrophic multiplexing. Figure 2.9 (a) shows the asymmetric reference and signal planewaves and the slanted grating vector. As the recording material is peristrophically rotated, the tip of the grating vector lifts off of the κ-sphere, causing the reconstruction to become Bragg mis-matched. When the recording material is peristrophically rotated by 180°, the grating vector is sticking way out of the κ-sphere as shown in Figure 2.9 (b). Since the grating vector knows no direction it can also be represented as shown in Figure 2.9 (c). Notice that the grating vector is again off the κ-sphere at the signal side. Therefore, depending on the thickness of the recording material, the flipped grating could be completely Bragg mis-matched. This means more holograms could be peristrophically recorded in the range of 180° to 360° without much cross-talk from the holograms stored in 0° to 180°.
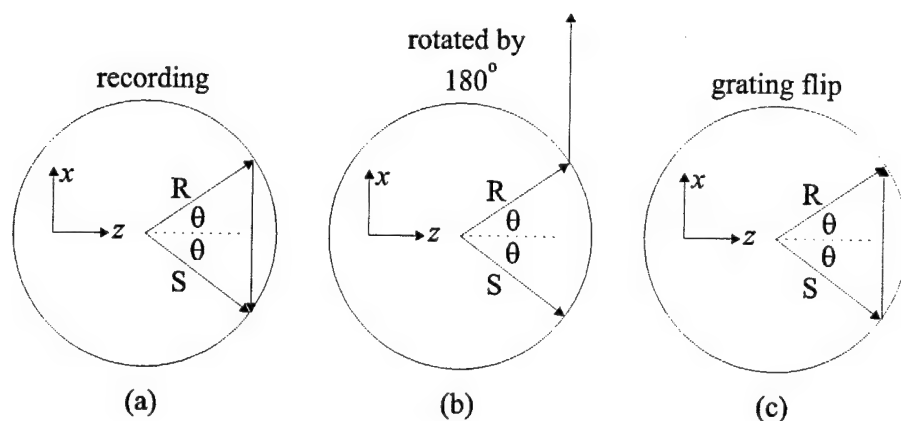
Figure 2.8 : The range of peristrophic multiplexing in symmetric planewave setup. (a) The recording geometry. (b) Peristrophic rotation of 180°. (c) At 180° rotation, the grating vector is once again on the κ-sphere.
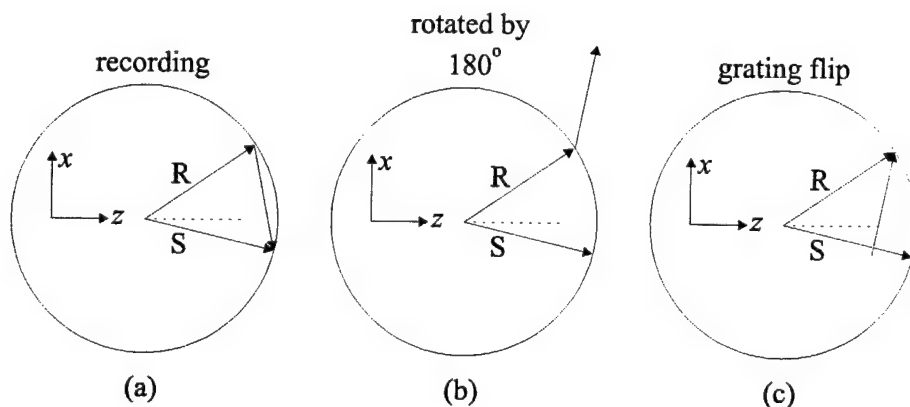

Figure 2.9 : The range of peristrophic multiplexing in asymmetric planewave setup. (a) The recording geometry. (b) Peristrophic rotation of 180°. (c) At 180° rotation, the grating vector is off of the κ-sphere regardless how it is anchored.

There is one more interesting recording geometry shown in Figure 2.10 (a). The reference beam is on axis and the signal beam is off axis with respect to the surface normal of the recording material. As the recording material is peristrophically rotated, the tip of the grating vector remains on the κ-sphere throughout the full 360° rotation (Figure 2.10 (b)). A full circle, with the reference beam at the center, can be traced out by rotating the recording material. This will have an important function later in peristrophic correlator systems.

recording

peristrophic
rotation

(a)   (b)
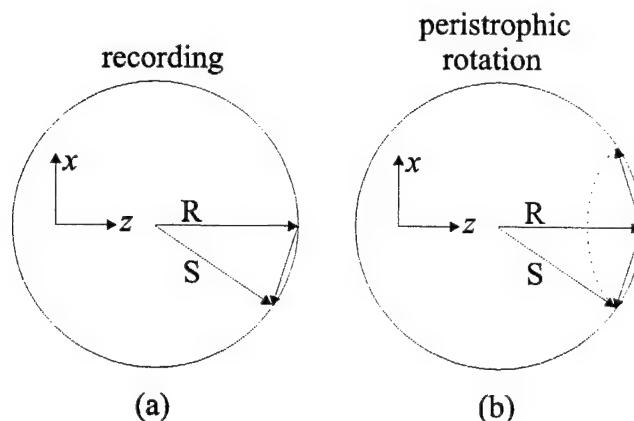
Figure 2.10 : A special case of the asymmetric planewave setup. (a) The reference beam is on axis while the signal beam is off. (b) For this special setup, the grating vector remains Bragg match throughout the entire range of peristrophic rotation.

So far we have talked about rotating the recording material as the means of achieving peristrophic multiplexing. The same effect can be achieved by rotating the reference beam around the surface normal of the recording material while keeping the hologram area illuminated. In most instances, it is much easier to just rotate the recording material.

2.3.3. Experimental Results

The first peristrophic multiplexing experiment was done in the Fresnel plane with the setup shown in Figure 2.11. We were unsure which selectivity (Fourier plane or image plane) was more dominant in the Fresnel plane, so both filtering mechanisms were incorporated in the setup. A random pixel data mask was used as the input image and we rotated the data mask in-plane (in the plane of the data mask) between holograms so that the reconstructed holograms could be distinguished. The data mask was imaged to the detector array with a pair of 4-f imaging systems. The recording material, DuPont's photopolymer, was placed significantly past the Fourier plane of the first 4-f imaging

system (Fresnel plane recording). This was done to allow the signal beam to become more uniformly distributed. A spatial filter with an aperture just large enough to pass the signal beam was positioned at the Fourier plane of the second 4-f system. This spatial filter served to block the undesired holograms for peristrophic multiplexing. Peristrophic multiplexing was achieved by rotating the photopolymer in-plane (in the plane of the photopolymer) with a rotational stage. Since the holograms were recorded in the Fresnel plane, the reconstructed images would both shift on the detector array and get blocked by the spatial filter as the recording material was rotated. The angle of the reference and signal beams, with respect to the surface normal of the photopolymer, were 32° and 29°, respectively. The intensity of the reference beam was 3.76 mW/cm^2 and the signal beam was 449 μW/cm^2.
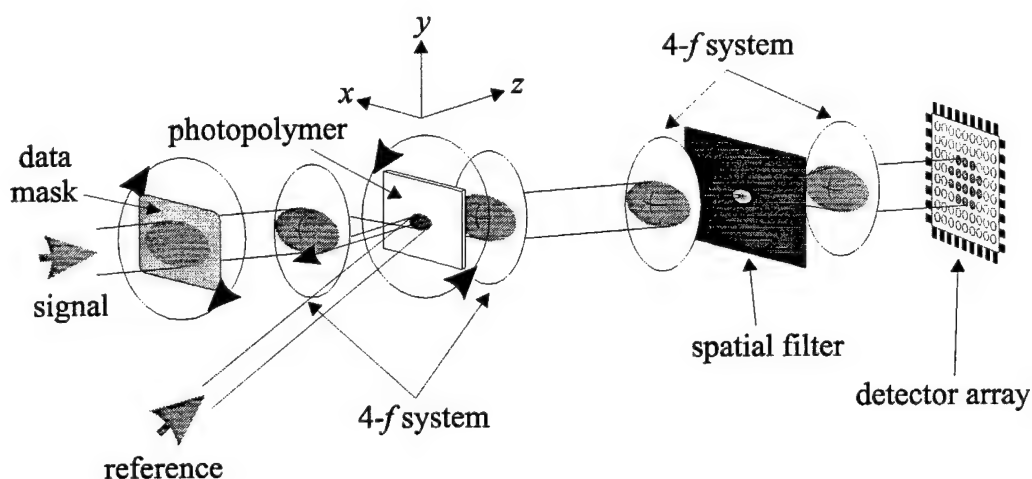
Figure 2.11 : The setup for the first peristrophic multiplexing experiment.

The peristrophic selectivity of this setup was experimentally determined to be 1.7° by recording one hologram in the 38 micron thick photopolymer and then slowly rotate the hologram until the reconstruction can no longer be seen by the detector array. This means that approximately 100 holograms could be stored using peristrophic

multiplexing from 0 to 180° in rotation. The full 360° in rotation cannot be used in this setup since the signal and reference beams were almost symmetric.

To demonstrate that this could indeed be done, 100 holograms were peristrophically multiplexed in DuPont's HRF-150-38, 38 μm thick photopolymer. The peristrophic separation between holograms was 1.7° and the data mask was rotated by -2° for each hologram. The exposure time per hologram was 0.12 seconds and there was a two second delay between holograms to allow the rotational stages to settle. Figure 2.12 shows some of the reconstructed holograms. The quality of the initial 70 or so holograms was very good. However, later holograms developed dark spots near the center of the image. These dark spots were caused by the depletion of the photopolymer's dynamic range. The center of the hologram is affected first since this is the part that gets exposed the most (the edges of the hologram is spread out over a larger area due to the rotation). Figure 2.13 shows the diffraction efficiency of all 100 stored holograms.

1st hologram

25th hologram

75th hologram

100th hologram
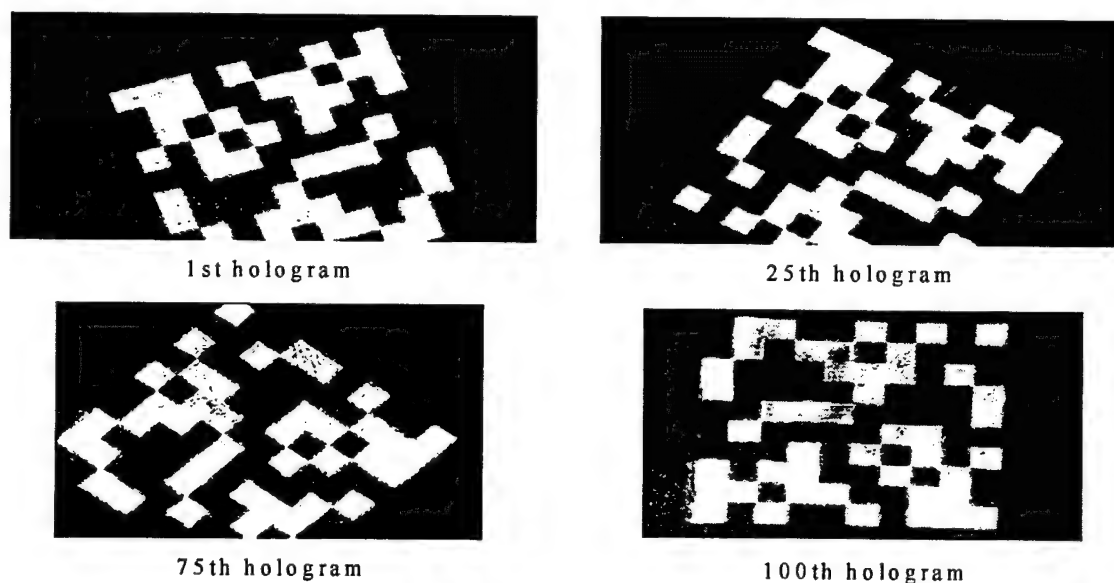
Figure 2.12 : Reconstructions from 100 peristrophically multiplexed holograms.
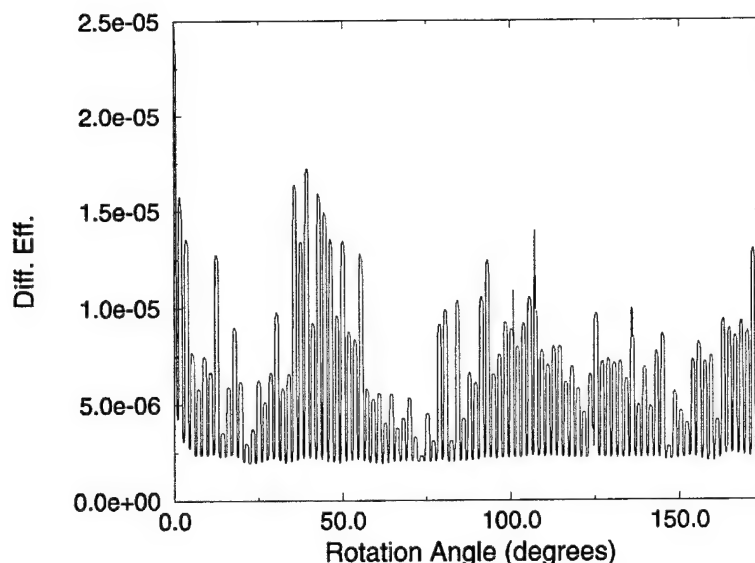
Figure 2.13 : Diffraction efficiency of 100 peristrophically multiplexed holograms (HRF-150-38).

The approximately sinusoidal variation in diffraction efficiency shown in Figure 2.13 is mostly due to the birefringent Mylar cover sheet on the photopolymer. For this experiment, the photopolymer was not laminated onto a glass substrate but was merely taped on, with both sides still covered with Mylar sheets. The weak birefringence of the Mylar cover sheets depolarizes the incident beams, causing the holograms not to record as strongly for certain peristrophic angles. We will come back and visit this issue later in Chapter 2.4.

Peristrophic multiplexing can be combined with angle multiplexing to further increase the number of holograms stored in a given volume. For the setup shown in Figure 2.14, a second rotational stage has been added to rotate the photopolymer around the y-axis. To the stored holograms, this rotation appears as a change in the reference beam angle and causes a Bragg mis-match effect very much like angle multiplexing.
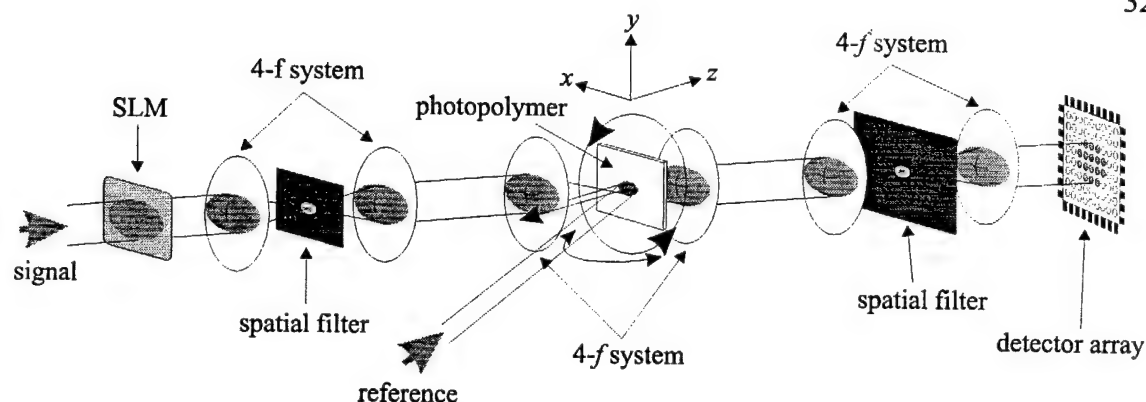
Figure 2.14 : Holographic data storage setup employing both angle and peristrophic multiplexing.

One other significant improvement in this setup is the use of a liquid-crystal type SLM for input. This allows us to store different images in each page of hologram. The extra 4-f system following the SLM is required to filter out the higher orders. Since the SLM doesn't have perfect fill factor, a grating diffraction pattern is produced under laser illumination. Only the zero[th] order is required so all higher orders were blocked by a spatial filter placed at the Fourier plane of the first 4-f system.

In this experiment, a sequence of cartoon frames were stored as holograms by using angle and peristrophic multiplexing. Each frame was numbered at the lower left corner according to the sequence in which they were stored. The reference and signal beams were initially incident at $\pm 30°$ with respect to the recording material's surface normal. The reference beam intensity was 1.1 mW/cm^2 and the signal beam had 300 μW illuminating an area of about 1cm by 0.5 cm on the photopolymer. Once again the recording material was placed a significant distance away from the focal point for better intensity distribution. A set of 5 angle multiplexed holograms were stored with 3° separation (rotation around the y-axis) and the photopolymer was rotated peristrophically

by 3° between each set of angle multiplexed holograms (using DuPont's HRF-150-38). A total of 59 sets of angle multiplexed holograms were peristrophically stored, giving a total of 295 multiplexed holograms. The initial exposure time was 0.11 seconds per hologram. However, starting with hologram number 26, each hologram was exposed for 0.005 seconds longer than the previous hologram to correct for the decrease in sensitivity. There was a 1.5 second delay between holograms to allow the rotational stages to settle. Figure 2.15 shows four reconstructions from the 295 holograms stored. The diffraction efficiency of all 295 holograms is shown in Figure 2.16.

1st Hologram

103rd Hologram

219th Hologram

295th Hologram

Figure 2.15 : Reconstructions from 295 angle and peristrophically multiplexed holograms.

The average diffraction efficiency of the stored holograms was $\sim 4 \times 10^{-6}$ and the variations are primarily due to the Mylar cover sheet and the intensity variation of the cartoon frames. The quality of the reconstructed images appeared grainy due to the low diffraction efficiency of the stored holograms. The scattering diffraction efficiency of the

HRF-150-38 photopolymer is also in the range of 10^{-6}. Therefore the signal-to-noise ratio (SNR) of the stored holograms was poor.
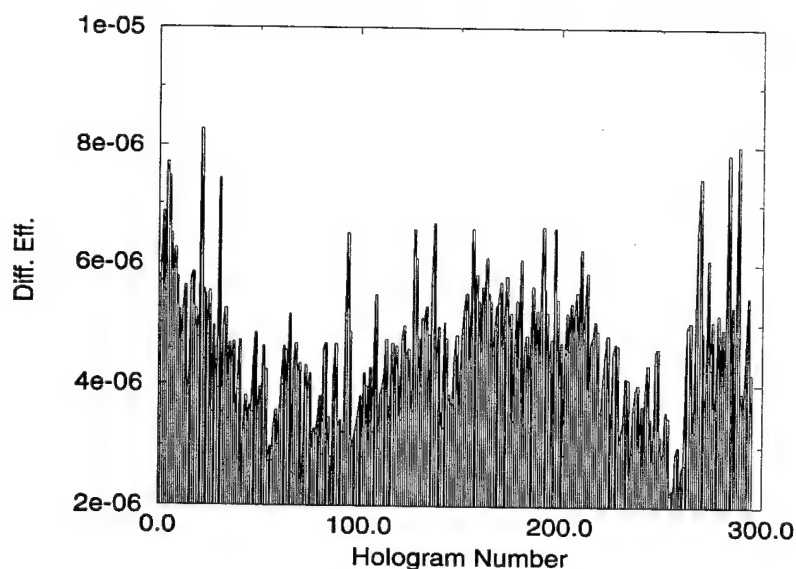

Figure 2.16 : Diffraction efficiency of 295 angle and peristrophically multiplexed holograms (HRF-150-38).

Previously, the number of holograms that could be stored in DuPont's HRF-150-38 (38 μm thick) photopolymer was system limited to 10. While the diffraction efficiency of the stored holograms was fairly high at 10^{-3}, there was no way to store more holograms in the same volume. By combining peristrophic with angle multiplexing, the number of holograms that could be stored in DuPont's photopolymer has shifted from system limited to material limited. The photopolymer runs out of dynamic range before more holograms could be stored. The focus of the next section is on how to utilize the limited dynamic range of the photopolymer more efficiently.

2.4 Photopolymer Characterization

2.4.1 Introduction

There have been several papers published on the characterization of the DuPont photopolymer [22-25]. What we are primarily interested in is how to best use the available dynamic range of the photopolymer for hologram multiplexing. Peristrophic multiplexing makes it possible to multiplex many holograms in thin films. As the number of holograms stored in these films increase, the optimal utilization of the available dynamic range becomes particularly important (since the diffraction efficiency scales as $1/M^2$, where M is the number of holograms multiplexed [26]). In this section we characterize the photopolymer by measuring the cumulative grating strength of the stored holograms as a function of exposure energy (the recording curve). By modeling the recording curve, we have a method for recording equal strength holograms that utilizes all of the available dynamic range of the recording material. This method can also be generalized to record unequal strength holograms or to use only a portion of the available dynamic range. Furthermore, no assumptions about the physics of the recording material are made and the method takes into account all experimental parameters including incident angles, incident intensities, and any special behavior of the recording material. We first demonstrate this method by recording 50 equal diffraction efficiency plane wave holograms using DuPont's HRF-150-38 photopolymer. Then, by recording different numbers of holograms, we measured the dependence of the diffraction efficiency on the number of holograms stored for the 38 and 100 micron thick photopolymers. Finally, 1,000 holograms of a random bit pattern were multiplexed in a 100 micron thick photopolymer using a recording schedule and the average diffraction efficiency was compared with the theoretical prediction.

2.4.2 Recording Curve

Figure 2.17 shows a planewave hologram setup using peristrophic multiplexing. The signal and reference planewaves both had an incident angle of 30° (outside) with respect to the surface normal of the photopolymer. The intensity of each beam was 250 $\mu W/cm^2$ and the wavelength was 488 nm. Shutters were used to block the beams in-between hologram recording and to block the signal beam during reconstruction. The photopolymer was mounted on a rotational stage for peristrophic multiplexing. After each exposure, the recording material was rotated in-plane by the rotational stage. This rotation causes the reconstruction from the stored hologram to deflect in a different direction, allowing for another hologram to be stored at the same location. Reconstructed holograms were Fourier transformed by a lens and then filtered by a small aperture placed at the Fourier plane. This spatial filter allows one reconstruction to pass through while blocking all the rest. Different holograms can be read-out by rotating the recording material so that its reconstruction is aligned with the spatial filter. The diffracted power of the reconstructed hologram is measured by a silicon detector placed after the spatial filter.

As stated in Section 2.3.3, how the photopolymer is prepared for recording affects the strength of peristrophically multiplexed holograms. Normally, the photopolymer sheet is cut into small 1" × 1" patches. The Mylar on one side is then removed, exposing the raw photopolymer. The photopolymer is naturally tacky (sticky) so it is very easy to laminate to a glass slide. However, it is usually very difficult to remove the Mylar on both sides of the photopolymer (unless the photopolymer has been exposed). Therefore,

the ready to use photopolymer slides usually come as glass-photopolymer-Mylar sandwiches.

Using the setup shown in Figure 2.17, 50 peristrophic holograms were multiplexed in the 38 micron thick photopolymer with the Mylar side facing the incident beams. The photopolymer was rotated by 2° between holograms and the signal and reference planewaves both had an intensity of 250 μW/cm^2. The exposure time per hologram was 2.5 seconds, giving each hologram an exposure energy of 1.25 mJ/cm^2. Figure 2.18 shows the diffraction efficiency of the 50 holograms as a function of the cumulative exposure energy that the photopolymer experienced.
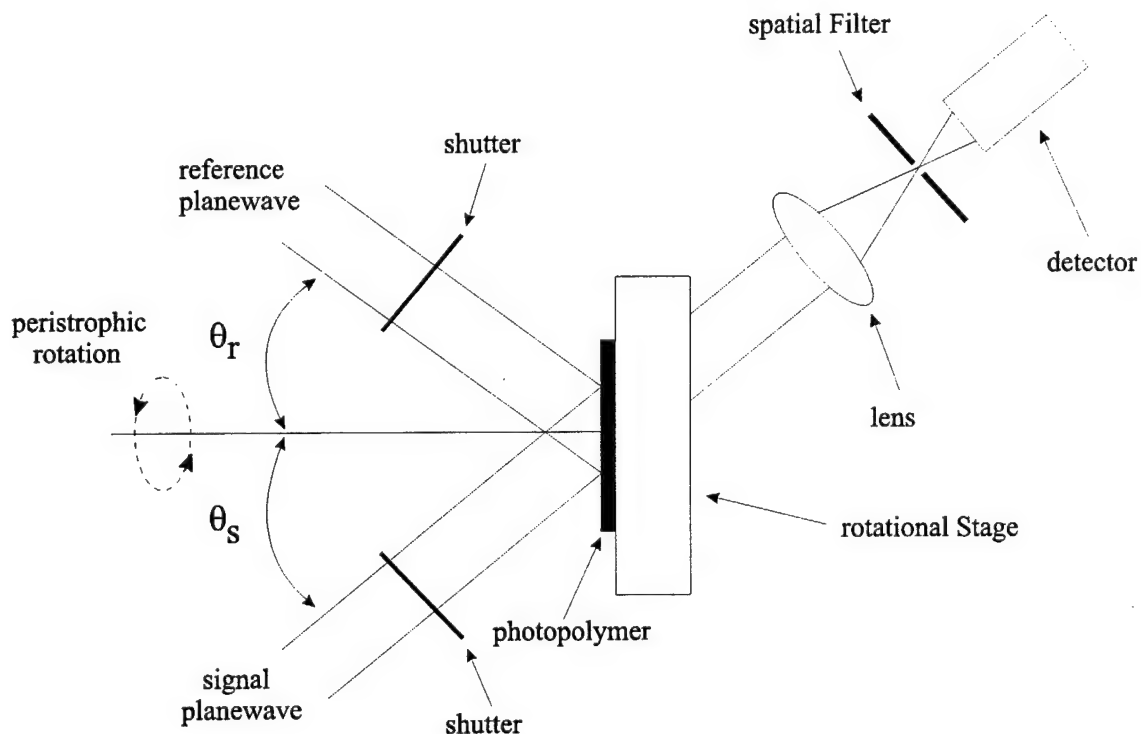
Figure 2.17 : A schematic diagram of the setup used to determine the recording curves.

One of the first things that is noticeable from Figure 2.18 is that the first 4 holograms are very weak. The 38 micron thick photopolymer is fairly insensitive until it has been exposed to about 6 mJ/cm^2 of incident energy. This much exposure energy is

required to sensitize (start the photoinitiated reaction) the photopolymer. The strength of the holograms then exhibits a sinusoidal variation until the material saturates. The sinusoidal variation is a strange behavior since each hologram was exposed to the same amount of energy. It is suspected that since the incident beams passed through the birefringent Mylar cover sheet first, the polarization of the incident beams were rotated during peristrophic multiplexing. Therefore, holograms at certain rotation angles recorded weaker, giving the sinusoidal shape. The diffraction efficiency of 50 peristrophically multiplexed holograms using the same experimental parameters, except that the glass substrate faced the incident beams first, is shown in Figure 2.19.
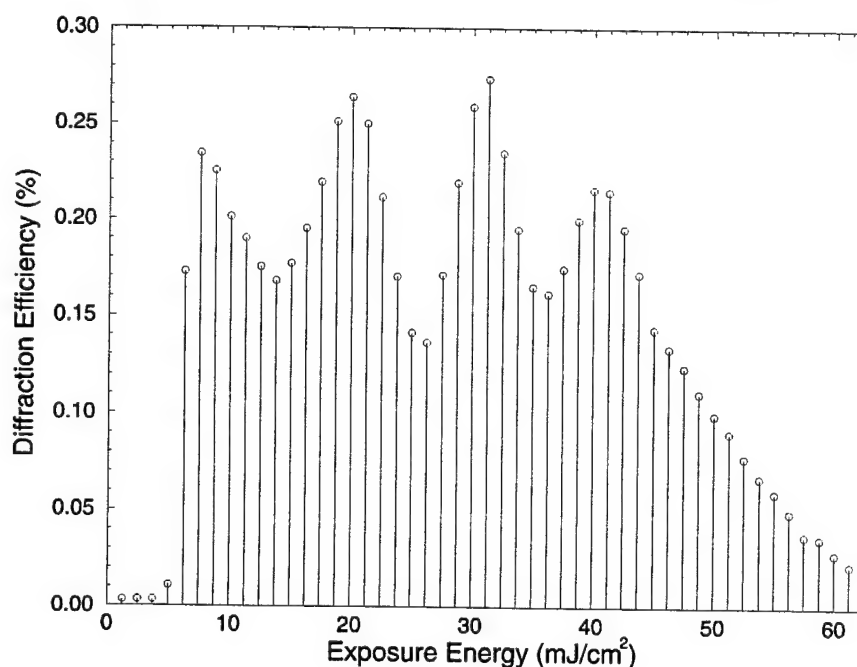
Figure 2.18 : Diffraction efficiency of 50 peristrophically multiplexed holograms with the Mylar side facing the incident beams.

Figure 2.19 seems to indicate that our suspicion is correct, that the Mylar cover sheet causes strange variations in the strength of peristrophically multiplexed holograms. By flipping the prepared photopolymer slide around to record holograms through the

glass face first, the Mylar film rotates the polarization only after the holograms have been recorded. This has no effect on the strength of the holograms nor the intensity of the reconstructions as the photopolymer is peristrophically rotated. To avoid any inconsistent recording behaviors, the glass slide should always face the incident beams first.
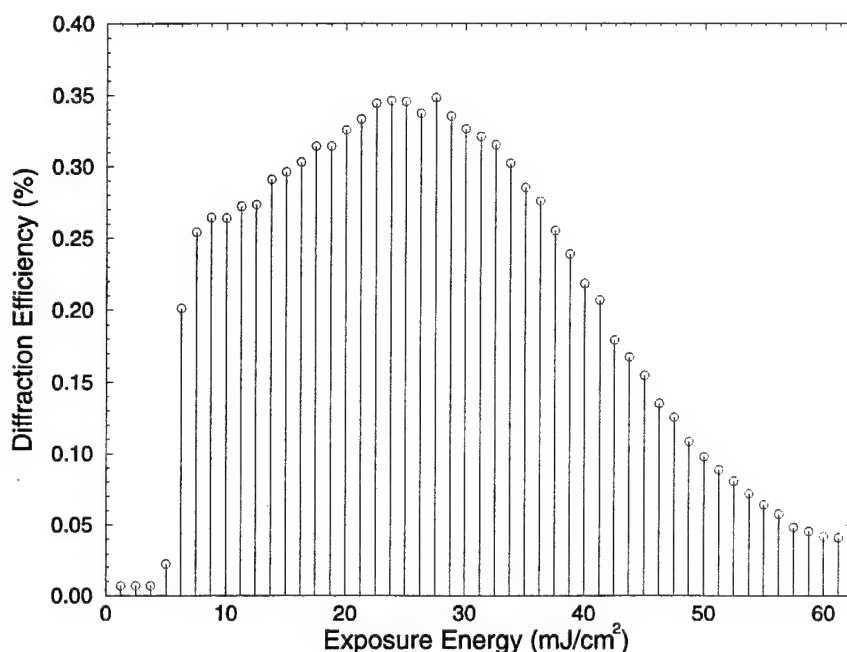

Figure 2.19 : Diffraction efficiency of 50 peristrophically multiplexed holograms with the glass side facing the incident beams.

In Figure 2.19, the sensitivity of the 38 micron thick photopolymer decreases once it has been exposed to approximately 30 mJ/cm^2 worth of energy. This is caused by the exhaustion of readily available monomers in the material. Subsequent exposures free up fewer monomers for grating formation. It is unclear from Figure 2.19 whether 62.5 mJ/cm^2 is the true material saturation point, beyond which no holograms will form regardless of exposure energy. Therefore, a larger exposure energy experiment is required to see the behavior of the photopolymer near saturation. The diffraction efficiency of 90 peristrophically (2° rotation between holograms) multiplexed planewave

holograms, recorded with a constant exposure schedule is shown in Figure 2.20. The holograms were recorded in the HRF-150-38 micron thick photopolymer with 1 mJ/cm^2 incident exposure energy per hologram (2 second exposure at .5 mW/cm^2 total incident intensity). The film exhibits the same insensitivity and then quasi-linear recording behavior until the material saturated. The hologram strength is not zero even after 90 mJ/cm^2 of exposure energy. However, it might not be worthwhile to pursue the remaining monomers with further exposure. At this point, a uniform UV exposure to cure the few remaining monomers is usually done.
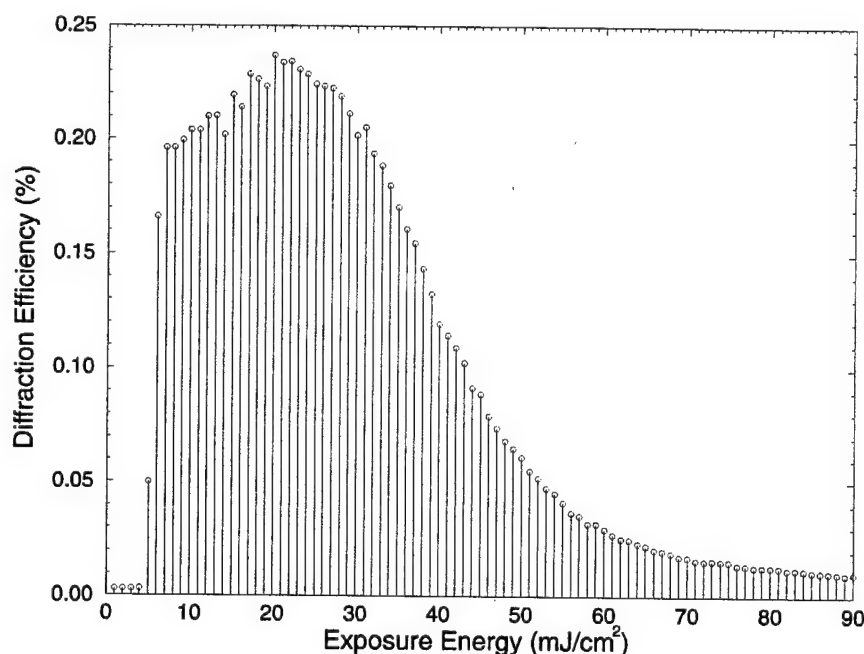
Figure 2.20 : The diffraction efficiency of 90 planewave holograms recorded with a constant 1mJ/cm^2 exposure per hologram.

The recording behavior of the photopolymer can be best characterized by a cumulative grating strength as a function of exposure energy curve (the recording curve). The diffracted power for each hologram is squared-rooted to obtain the grating strength and the grating strength of all the stored holograms are then summed with respect to the

cumulative exposure energy to generate the recording curve. The recording curve is important since it shows grating formation as a function of exposure energy. By taking the derivative of the recording curve, the grating formation rate as a function of exposure energy is obtained.

Personally, I believe recording curves generated using the above method reflect more accurately real applications where many holograms are multiplexed in the same volume. Most researchers generate a diffraction efficiency as a function of exposure energy curve in real-time by monitoring the formation of one hologram with a Bragg matched probe beam. While this accurately reflects the grating dynamics of one hologram, the curve usually cannot be applied where the formation of many holograms are concerned (due to "run-time" effects which will be discussed shortly).

The curves shown in Figure 2.21 were obtained by integrating the square-root of the diffracted power from peristrophically multiplexed holograms recorded with 1 mJ/cm^2, 1.5 mJ/cm^2, and 2 mJ/cm^2 constant exposure schedules, and 6 mJ/cm^2 pre-exposures. The origin of the exposure energy axis in Figure 2.21 has been reset to leave out the pre-exposure energy. From the figure, it can be seen that the cumulative grating strength grows quasi-linearly with exposure energy and then saturates. The final saturation grating amplitude for 1 mJ/cm^2, 1.5 mJ/cm^2, and 2 mJ/cm^2 exposure per hologram is different due to something called the "run-time" effect [16]. For the three different exposure energies per hologram, the total exposure energy and incident intensities were kept constant at 90 mJ/cm^2 and 250 $\mu W/cm^2$ per beam, respectively, while the number of holograms and exposure time for each of the three experiments were changed. After a hologram was written, the beams were blocked, the film was rotated for

peristrophic multiplexing, and 1.5 seconds was allowed to elapse before a new hologram was recorded so that the table could settle. During that 1.5 seconds, the material is losing dynamic range because once the photon-induced reaction is started, it continues even in the absence of light. This mechanism is the origin of the "run-time" effect. The 2 mJ/cm^2 per hologram curve wasted less dynamic range since a larger fraction of the "run-time" was spent on recording (4 second exposure per hologram for only 45 holograms). Therefore, the 2 mJ/cm^2 per hologram curve has the highest saturation grating amplitude.
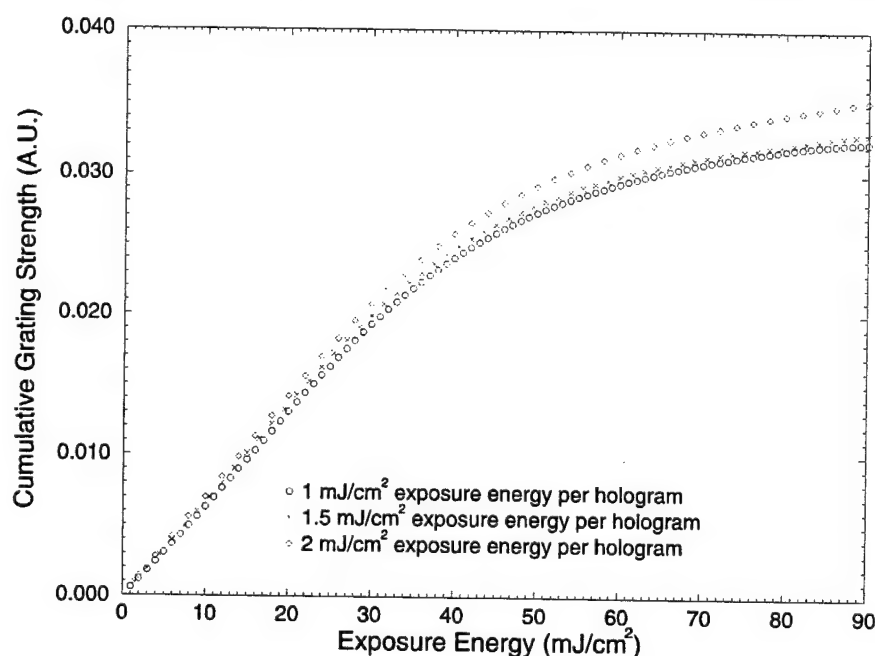
Figure 2.21 : The cumulative grating strength as a function of exposure energy for three different exposure energies per hologram.

The 1.5 mJ/cm^2 curve in Figure 2.21 shows only a slight increase in saturation grating amplitude over the 1 mJ/cm^2 curve, while the 2 mJ/cm^2 curve shows a more significant improvement. Theoretically, it should be the other way around since the reduction in total waiting time is greater going from 1 mJ/cm^2 to 1.5 mJ/cm^2 per hologram than from 1.5 mJ/cm^2 to 2 mJ/cm^2 per hologram. However, for the 1 mJ/cm^2

curve, 90 holograms were recorded and their diffracted powers square-rooted and integrated. Therefore, the background noise from scattering gets integrated 90 times for the 1 mJ/cm^2 curve as compared to only 45 times for the 2 mJ/cm^2 curve. This boosts the 1 mJ/cm^2 curve artificially closer to the 1.5 mJ/cm^2 curve.

The recording curves shown in Figure 2.21 are for peristrophic holograms recorded with a 488 nm argon laser. The same DuPont HRF-150 material can also be used with the compact and highly efficient 532 nm diode-pumped-solid-state (DPSS) lasers. It is known that the sensitivity of the HRF-150 photopolymer is reduced at the longer wavelengths due to a lower absorption coefficient. However it is not known if this reduction in sensitivity will cause the photopolymer to exhibit a smaller dynamic range. Figure 2.22 shows the recording curve for 90 peristrophically multiplexed holograms (2° separation) recorded at 532 nm. The signal and reference planewaves were nearly symmetric with respect to the surface normal of the recording material (θ_r=29.5^o and θ_s=33°). The intensity of each beam was 1 mW/cm^2 and the holograms were recorded with a constant 1 second exposure. The settling time between holograms was 1.5 seconds.

The recording curve at 532 nm looks very similar to the recording curve at 488nm (Figure 2.21), except that the exposure energy axis is considerably lengthened. It now requires an exposure energy of 16 mJ/cm^2 to sensitize the photopolymer at 532 nm and material saturation doesn't occur until approximately 150 mJ/cm^2. However, the cumulative grating strength at saturation is about the same for 532 and 488 nm exposures. Therefore, the dynamic range of the HRF-150 photopolymer remains unchanged but the

photo-speed of the photopolymer at 532 nm is much reduced when compared with recording at 488 nm.
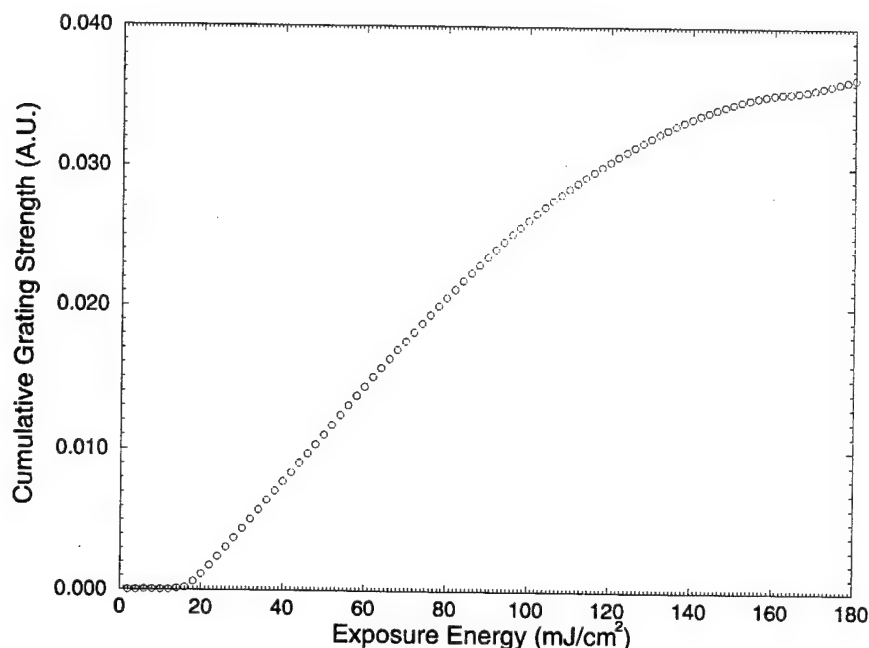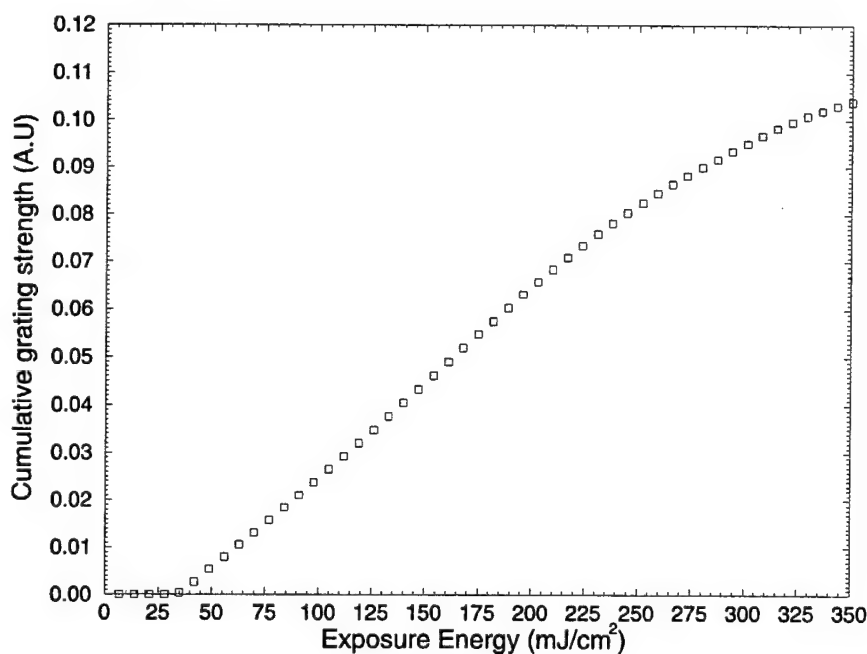

Figure 2.22 : The cumulative grating strength as a function of exposure energy for 90 peristrophic holograms recorded with a 532 nm DPSS laser.

Figure 2.23 shows the recording curve for 50 peristrophic holograms recorded with a 532 nm DPSS laser, using the HRF-150-100 (100 micron thick) photopolymer. The signal and reference planewaves were symmetric with respect to the photopolymer's surface normal ($\theta_r = 30^\circ = \theta_s$). The intensity of the incident beams were 250 μW/cm^2 a piece and each hologram was exposed to 7 mJ/cm^2 worth of energy (14 seconds exposure per hologram). The separation between peristrophic holograms was 2° and a 1 second delay allowed the table to settle.

Approximately 36 mJ/cm^2 worth of exposure energy is required to sensitize the 100 micron thick photopolymer at 532 nm. The saturation point of the 100 micron thick photopolymer is around 350 mJ/cm^2, about 2.5 times higher than the 38 micron thick

photopolymer. Furthermore, the final cumulative grating strength at saturation is approximately three times higher than the 38 micron thick photopolymer. Due to the increased thickness, the amount of raw photopolymer material for grating formation is more than doubled. Therefore, to completely exhaust the available monomers, it takes a larger amount of exposure energy. Since there are more monomers for grating formation, the final cumulative grating strength is also higher. This means the 100 micron thick photopolymer has more dynamic range when compared to the 38 micron version. We will come back to the dynamic range issue later in Chapter 2.4.4 when the *M/#* for these materials is determined.

Figure 2.23 : The cumulative grating strength as a function of exposure energy for 50 peristrophic holograms recorded with a 532 nm DPSS laser using HRF-150-100 photopolymer.

2.4.3 Exposure Schedule

In Figure 2.20, each hologram was exposed to a constant 1 mJ/cm^2 exposure but their diffraction efficiencies were not the same. For most applications, it is highly desirable to have equal strength holograms. The photopolymer requires a small pre-exposure to sensitize the material. Once sensitized, the photopolymer exhibits a short quasi-linear recording period and then the diffraction efficiency drops as the material saturates. Near equal strength holograms can be recorded in the short quasi-linear region with constant energy exposures. However, that means a significant portion of the photopolymer's dynamic range would not be used.

Utilizing the entire dynamic range of the photopolymer to store equal strength holograms is the subject of this section. We introduce an iterative method for determining the exposure schedule for multiplexing holograms in saturable recording materials. This method is designed to share all or part of the available dynamic range of the recording material among the holograms to be multiplexed. The photopolymer is first pre-exposed to sensitize the material (by illuminating it with just the reference beam). Then a recording schedule that compensates for the loss in sensitivity as the material saturates is used to equalize the holograms.

In order to derive an exposure schedule that equalizes the strength of the holograms for the HRF-150-38 photopolymer, we need to fit the recording curves of Figure 2.21 to an analytical expression. The 1 mJ/cm^2 per hologram curve was fitted to a sixth order polynomial of the following form:

$$A(E) = a_0 + a_1 E + a_2 E^2 + a_3 E^3 + a_4 E^4 + a_5 E^5 + a_6 E^6 \qquad (2.7)$$

where $A(E)$ is the cumulative grating strength, E is the cumulative exposure energy, and a_0-a_6 are constants from the fit. A sixth order polynomial was used because it was found

standard deviation is only .006 (in % diffraction efficiency). The above procedure can be

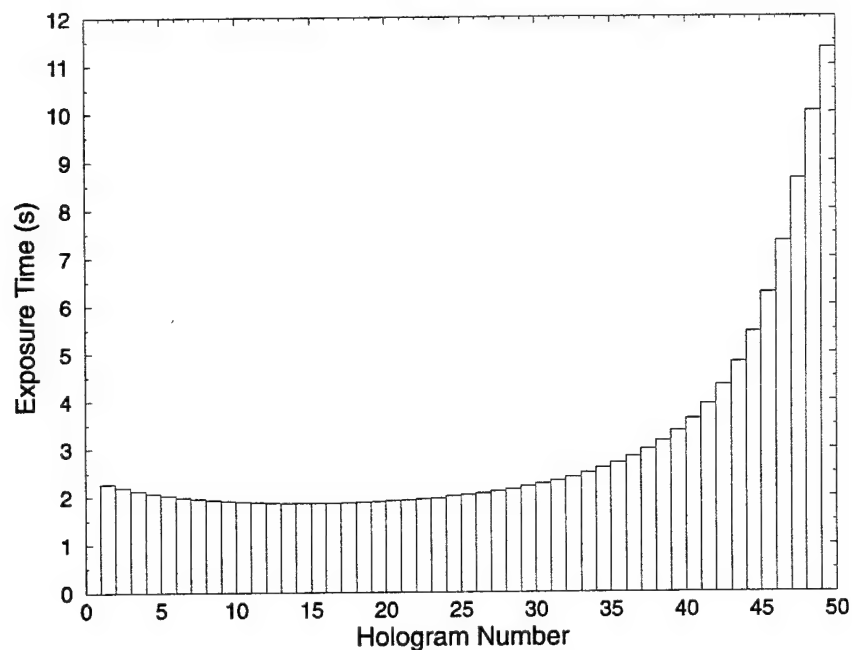repeated until the desired uniformity is reached.

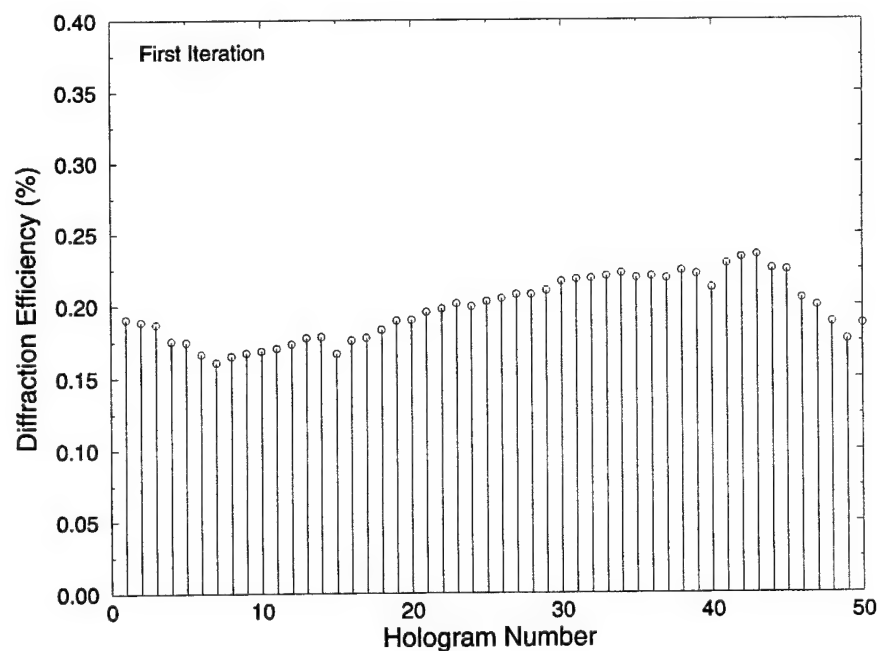Figure 2.24 : A first iteration exposure schedule for 50 holograms.

Figure 2.25 : The diffraction efficiency as a function of hologram number for holograms recorded with the first iteration exposure schedule.

fitting Figure 2.21. Figure 2.24 shows the resulting exposure schedule. Notice that the exposure time per hologram increases as the film becomes saturated. The diffraction efficiencies of the 50 holograms that were recorded with this schedule is shown in Figure 2.25. The grating strength of these 50 holograms are more uniformly distributed compared to the ones in Figure 2.20 but still not equal. This first iteration does not give us perfectly uniform holograms due to the fact that some of the recording parameters have been changed. The exposure time per hologram is no longer a constant and we are now recording 50 holograms instead of the 90 used to generate the recording curve. Therefore, these effects have to be remodeled. This can be done by iterating the previous procedure. Specifically, we integrate the square-root of the diffracted power from the 50 holograms recorded with the schedule and plot it as a function of exposure energy. We fit this curve to a sixth order polynomial, and then generate a new schedule using Eq. 2.9. Figure 2.26 shows the resulting cumulative grating strength as a function of exposure energy for the scheduled recording. Notice that the saturation grating strength for the scheduled recording is higher than the constant time recording. This is due to the fact that with only 50 holograms stored, the "run-time" effect is less. Another possibility is that scheduled recording uses the dynamic range of the recording material more efficiently though better managed exposures. When the photopolymer is sensitive, only a short exposure time is required. When the photopolymer is less sensitive, a longer exposure time activates the remaining monomers to achieved the same grating strength.

Figure 2.27 shows the diffraction efficiency of 50 holograms that were recorded with the second iteration schedule generated from Figure 2.26. From Figure 2.27 it can be seen that the grating strength is quite uniform among the 50 holograms and the

to have sufficient degrees of freedom to fit closely with Figure 2.21. Other more elegant expressions such as exponentials and hyperbolic tangents were also tried. However, they did not fit closely enough to produce good results.

By taking the derivative of Eq. 2.7, the grating strength growth rate as a function of exposure energy is obtained. Therefore, it is now straightforward to pre-determine the strength of each hologram by appropriating the exposure energy correctly. For example, equal strength holograms can be achieved by allocating the entire dynamic range of the photopolymer equally among the holograms. In that case, the desired exposure schedule becomes:

$$\frac{A_{sat}}{M} = \left.\frac{\partial A(E)}{\partial E}\right|_{E=\sum_{i=1}^{n-1} E_i} \times E_n \tag{2.8}$$

where A_{sat} is the saturation grating strength from Figure 2.21, M is the number of multiplexed holograms, E_i is the amount of energy the photopolymer received in order to record the ith hologram, and E_n is the amount of energy required to record the nth hologram. Each hologram is allocated $1/M^{th}$ of the photopolymer's dynamic range.

Using the same exposure intensity of 250 μW/cm^2 per beam (so our fit of Figure 2.21 remains valid), the exposure schedule in terms of exposure time becomes:

$$t_n = \frac{A_{sat}}{M\left(a_1 + 2a_2\sum_{i=1}^{n-1} E_i + 3a_3\left(\sum_{i=1}^{n-1} E_i\right)^2 + 4a_4\left(\sum_{i=1}^{n-1} E_i\right)^3 + 5a_5\left(\sum_{i=1}^{n-1} E_i\right)^4 + 6a_6\left(\sum_{i=1}^{n-1} E_i\right)^5\right)I} \tag{2.9}$$

where t_n is the exposure time of the nth hologram and I is the total incident intensity.

An exposure schedule for 50 holograms was calculated by the finite step method (step size = 0.001 mJ/cm^2) using Eq. 2.9 and the parameters a_1 through a_6 obtained from
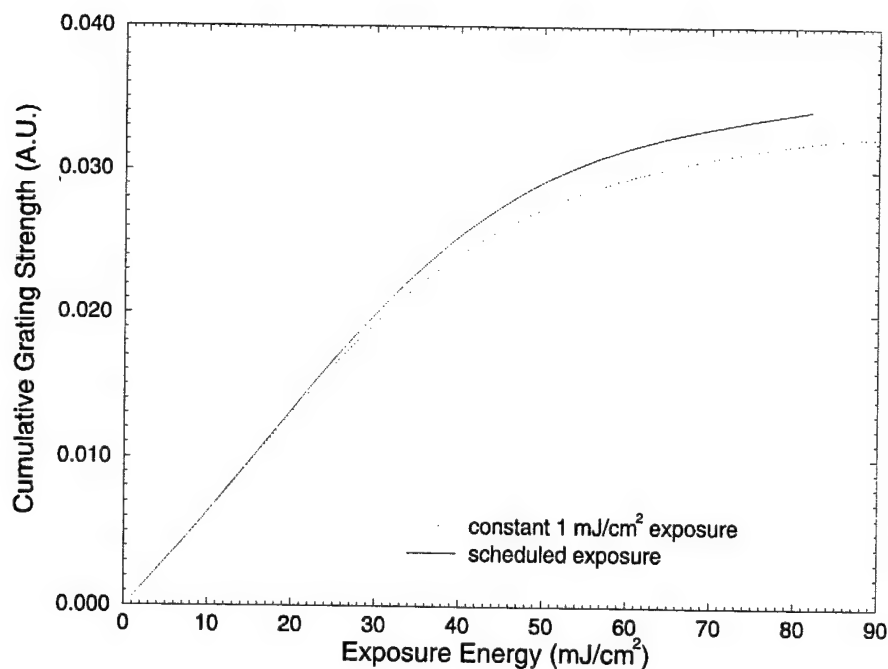
Figure 2.26 : Comparison between the cumulative grating strength of 50 holograms recorded with a schedule and 90 holograms recorded with a 1 mJ/cm^2 constant exposure.
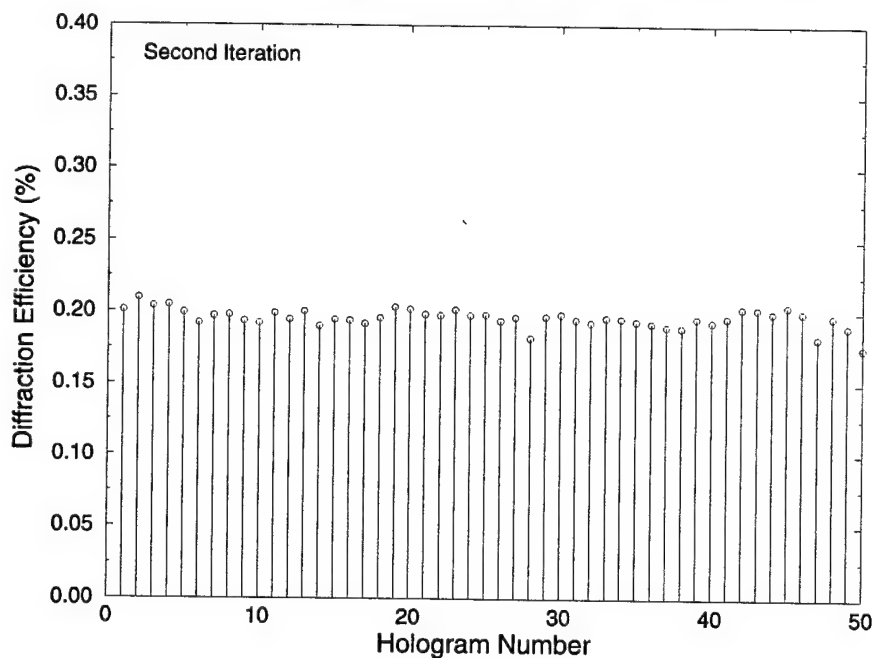

Figure 2.27 : The diffraction efficiency as a function of hologram number for holograms recorded with the second iteration exposure schedule.

The same scheme can also be applied to record equal strength holograms in the 100 micron thick photopolymer. Using the same setup as for the 38 micron thick

photopolymer, Figure 2.28 shows the exposure schedule for 50 peristrophic holograms after two iterations and Figure 2.29 shows the resulting holograms. Notice the longer exposure time and higher diffraction efficiency achieved with the 100 micron thick photopolymer when compared with Figures 2.24 and 2.27. This is due to the 100 micron thick photopolymer's larger dynamic range, A_{sat}, which is divided equally among the 50 multiplexed holograms. The exposure time can be decreased significantly by increasing the incident intensities.
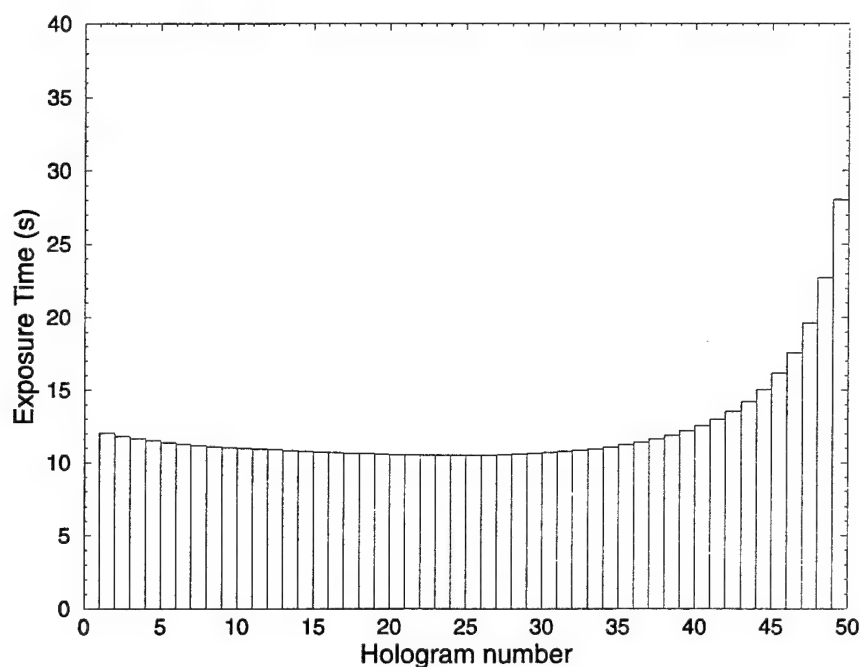

Figure 2.28 : Second iteration recording schedule for 50 peristrophic holograms in HRF-150-100 photopolymer.

Eq. 2.8 shows how uniform strength holograms can be multiplexed in the same volume by sharing the dynamic range of the recording material equally. Unequal strength holograms of any design can also be recorded by replacing A_{sat}/M in Eq. 2.8 by a function. For example, holograms with strengths that resemble a stairwell or a sawtooth can be made by allocating the dynamic range of the recording material appropriately as a

function of hologram number. The iterative nature of this method will converge to the correct exposure schedule. This simple method of determining the recording schedule can be used for any number of holograms and any recording setup. Moreover, since we did not make use of the physical properties of the recording medium in deriving the exposure schedule, this method for recording equalized holograms can be applied to any saturable recording medium.
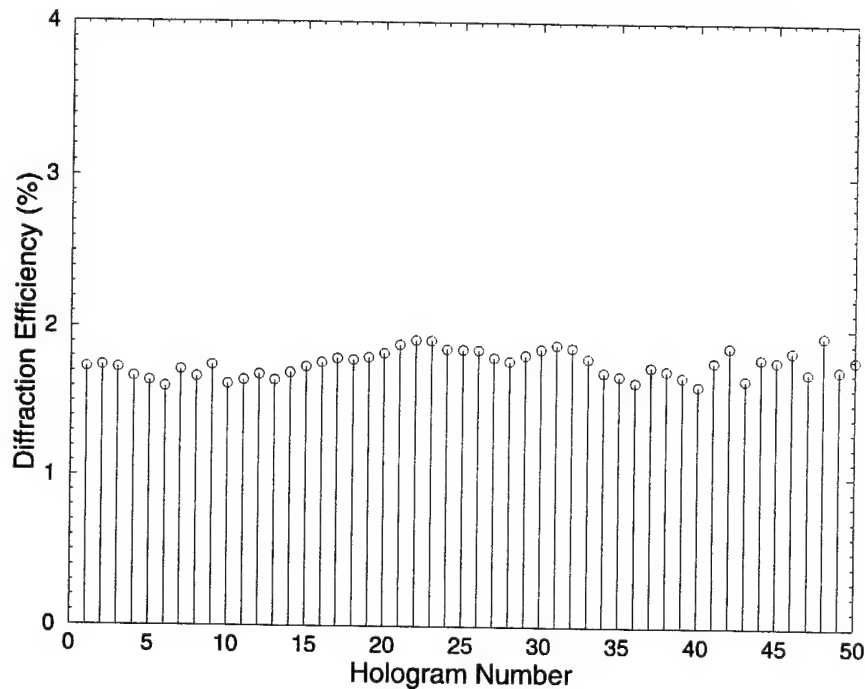

Figure 2.29 : The diffraction efficiency of 50 peristrophically multiplexed holograms using the second iteration schedule (HRF-150-100).

2.4.4 *M/#* Measurements

We can use the recording schedule to determine how the diffraction efficiency per hologram decreases as the number of multiplexed holograms increase. Previously, in the determination of the exposure schedule for equal strength holograms, each hologram was allocated a grating strength of A_{sat}/M. Since diffraction efficiency is proportional to the

grating strength squared, the diffraction efficiency scales as α/M^2, where α is a proportionality constant. To determine α, 10, 25, 50, 75 and, 90 peristrophic planewave holograms were recorded using schedules obtained with the above method for the HRF-150-38 and HRF-150-100 photopolymer films (for the 100 micron thick film, only 25, 50, 75 and 90 planewave holograms were recorded). The holograms were all equal strength and used the entire dynamic range of the recording material. The resulting diffraction efficiencies for these cases are plotted in Figure 2.30. From fitting Figure 2.30, the square-root of α was determined to be 2.2 for the 38 micron thick photopolymer and 6.5 for the 100 micron thick photopolymer. For photorefractives, the square-root of α is called the $M/\#$, and we shall also call the square-root of α the $M/\#$ for the photopolymer.
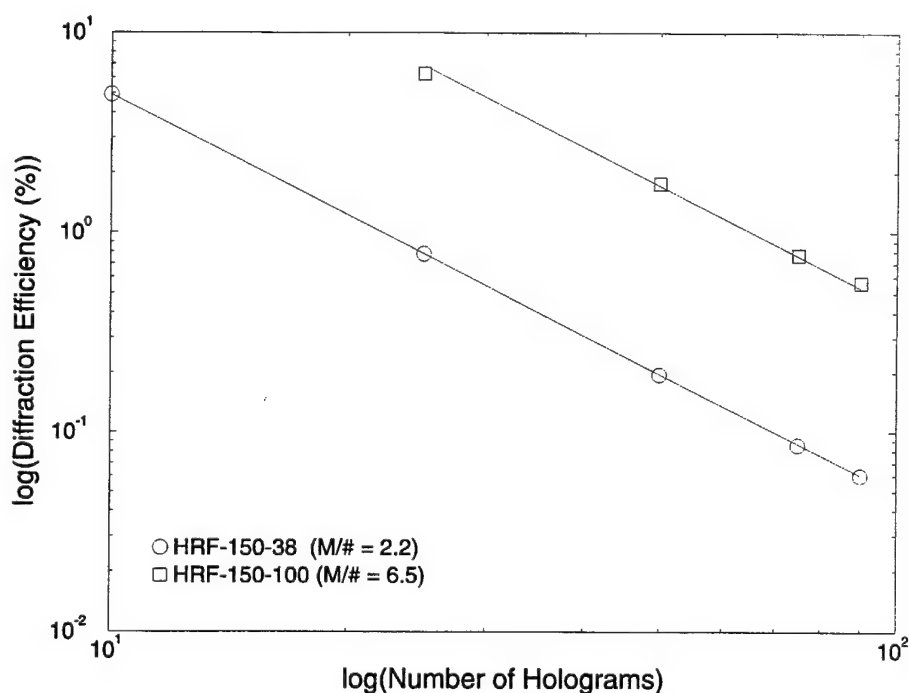

Figure 2.30 : The diffraction efficiency as a function of number of holograms multiplexed in the same volume.

The M/# of a recording material varies depending on experimental conditions such as material thickness, absorption, recording geometry, and the ratio of reference to signal beam intensity. After the M/# of a material is measured for a particular setup, it can be used to predict the diffraction efficiency obtainable for the number of holograms to be multiplexed. The relationship is:

$$\eta = \left(\frac{M/\#}{M}\right)^2 \qquad\qquad (2.10)$$

where η is the diffraction efficiency and M is the number of holograms to be multiplexed.

For example, if we were to multiplex 1,000 holograms in the HRF-150-100 photopolymer, the theory predicts a diffraction efficiency of $(6.5/1,000)^2 \cong 4 \times 10^{-5}$ per hologram. To find out if the theory is any good, an experiment was devised to store 1,000 image plane holograms of a random 100 × 100 pixel pattern in the HRF-150-100 photopolymer.

2.4.5 1,000 Hologram Experiment

For this experiment, the size of each pixel in the data mask was 100 μm by 100 μm and it was made from a photographic glass plate. To multiplex 1,000 holograms in the 100 μm thick photopolymer, peristrophic and angle multiplexing were combined. Figure 2.31 shows the experimental setup. The reference and signal beams were initially incident at ±30° from the film's normal. Angle multiplexing was achieved by rotating the recording material in the plane of interaction. The reference and signal beam intensities were 1 mW/cm² per beam. 1,000 holograms were multiplexed at the same location by recording 100 peristrophic holograms at each of the 10 angle multiplexing positions. The

film was rotated in the plane of the photopolymer by 1.8° for peristrophic multiplexing. Peristrophic image plane theory predicts that a rotation greater than .6° is required. Angular positions were separated by 1.5°. There was a 1 second delay between holograms to allow the rotational stages to settle completely. Figure 2.32 shows the angular selectivity of this setup.
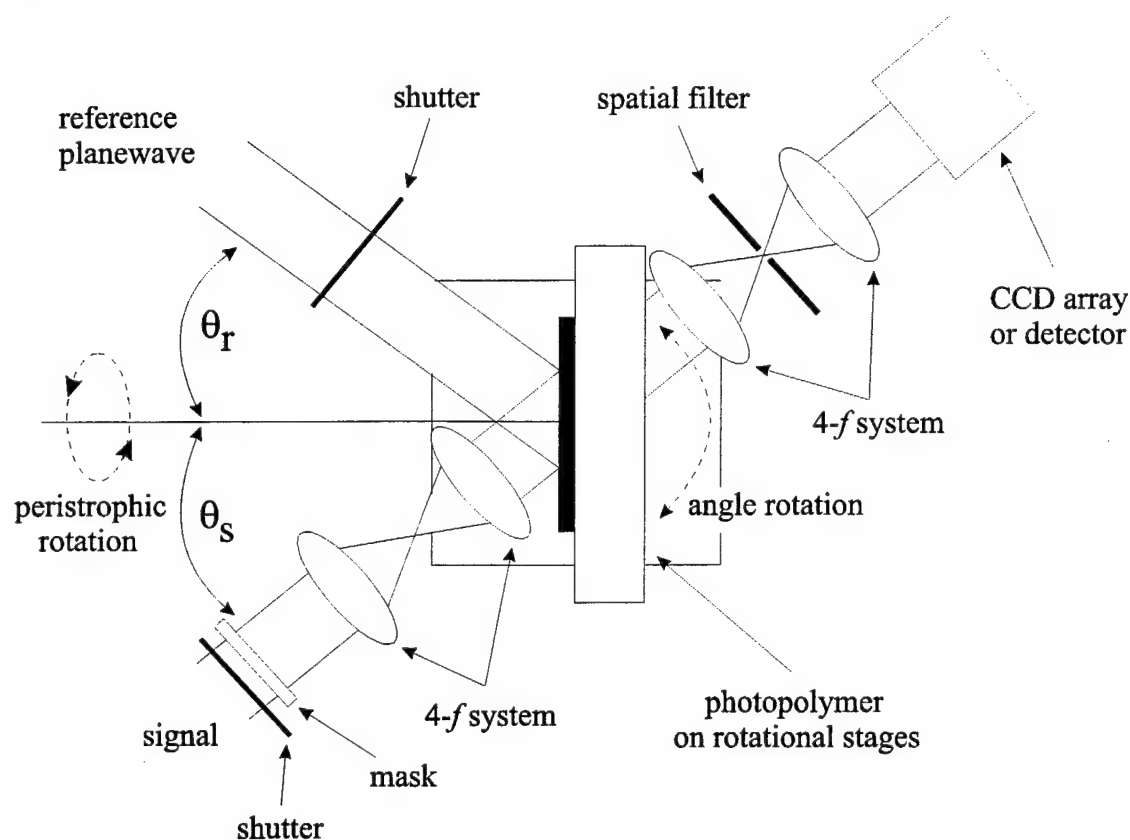

Figure 2.31 : A schematic diagram of the setup used to record 1,000 holograms in HRF-150-100 photopolymer.

An average diffraction efficiency of ~4x10^{-5} (corrected for the half-on, half-off nature of the binary amplitude mask) was achieved with the 1,000 stored holograms using six iterations of the recording schedule. The measured diffraction efficiency matched very well with the $M/\#$ theory discussed in the previous section. Figure 2.33 shows the resulting diffraction efficiency of the 1,000 holograms. Some of the sharp non-linearity

in the curve is due to stage instability during recording. The rotational stages used in this experiment for multiplexing are servo units and they sometimes twitch unpredictably.
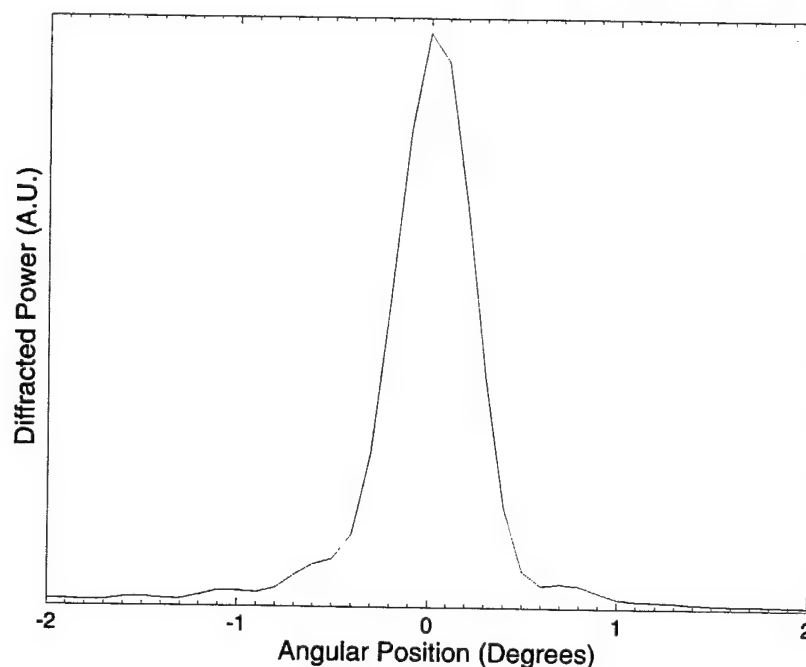

Figure 2.32 : Angular selectivity of the 1,000 hologram setup using the HRF-150-100 photopolymer.

Other causes for the non-uniformity in the measured diffraction efficiency are photopolymer shrinkage during recording and laser fluctuations. Shrinkage causes the holograms to Bragg-match at a slightly different reference beam angle than that used to record the hologram. Therefore, to measure the true diffraction efficiency of the holograms, it is necessary to find the maximum diffracted power by de-tuning the reference beam angle for each hologram. One way to achieve this is to scan the reference beam angle in small steps during reconstruction and note the maximum diffracted power for each hologram. However, this would require us to take power measurements at an order of magnitude higher than 1,000 points. At 4 seconds per point (to allow the power meter to settle), the entire measurement could run into days. Therefore, we opted to

measure the diffracted power at the same reference beam angle as used for recording. The shrinkage effect was not noticeable in the other experiments discussed previously because the signal and the reference beams were both fixed at 30° with respect to the surface normal and no angle multiplexed holograms were recorded. For symmetric recording ($\theta_s = \theta_r$), shrinkage does not change the Bragg condition (more on shrinkage in Chapter 3).
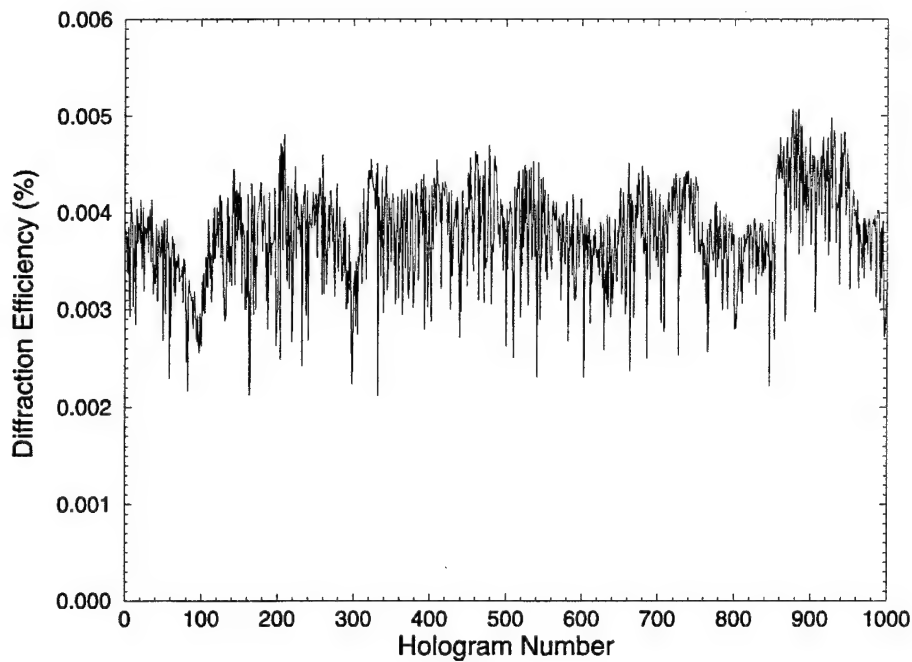

Figure 2.33 : The diffraction efficiency as a function of hologram number for the 1,000 hologram experiment.
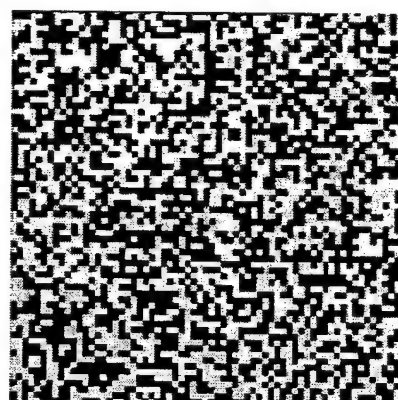
Figure 2.34 shows the original image and 3 reconstructed holograms. The signal-to-noise ratio of the original image and the reconstructions can be calculated using [27]:

$$SNR = \frac{m_2 - m_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} \qquad (2.11)$$

where m_1 and σ_1 are the mean and standard deviation for the dark areas. Similarly, m_2 and σ_2 are the mean and standard deviation for the bright areas. The average SNR for

some sampled reconstructions from the 1,000 holograms experiment is approximately 7.5 while the SNR of the original image is 12. Assuming gaussian noise statistics, this gives a bit-error-rate of better than 10^{-12} for the reconstructions.
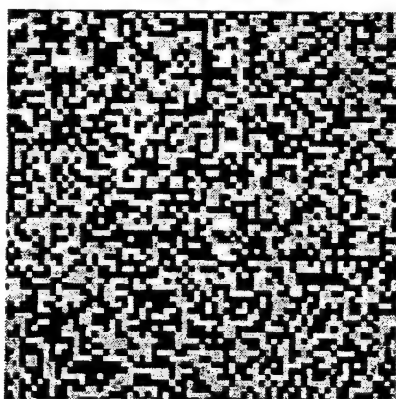
Figure 2.35 shows a line cross-section of the original image and 3 reconstructed holograms. The horizontal axis is the CCD pixel number while the vertical axis is the CCD pixel value. The CCD over-samples the data mask pixels since its pixels are much smaller than 100 μm. From Figure 2.35 it can be seen that even for the 1,000th hologram, the dark data mask pixels are clear distinguishable from the bright data mask pixels. For example, a pixel value threshold of 35 could be used to classify the data mask pixels for all the reconstructions.
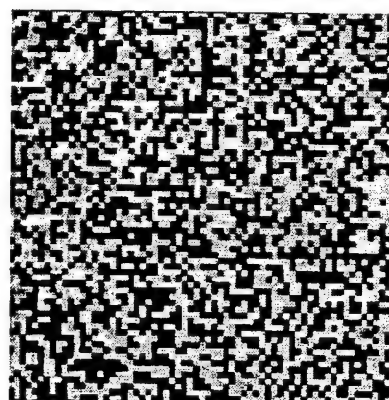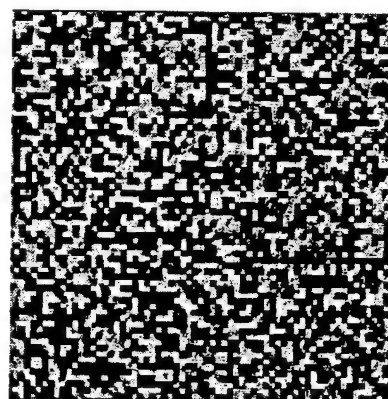

Original Image

1st Hologram

500th Hologram

1,000th Hologram

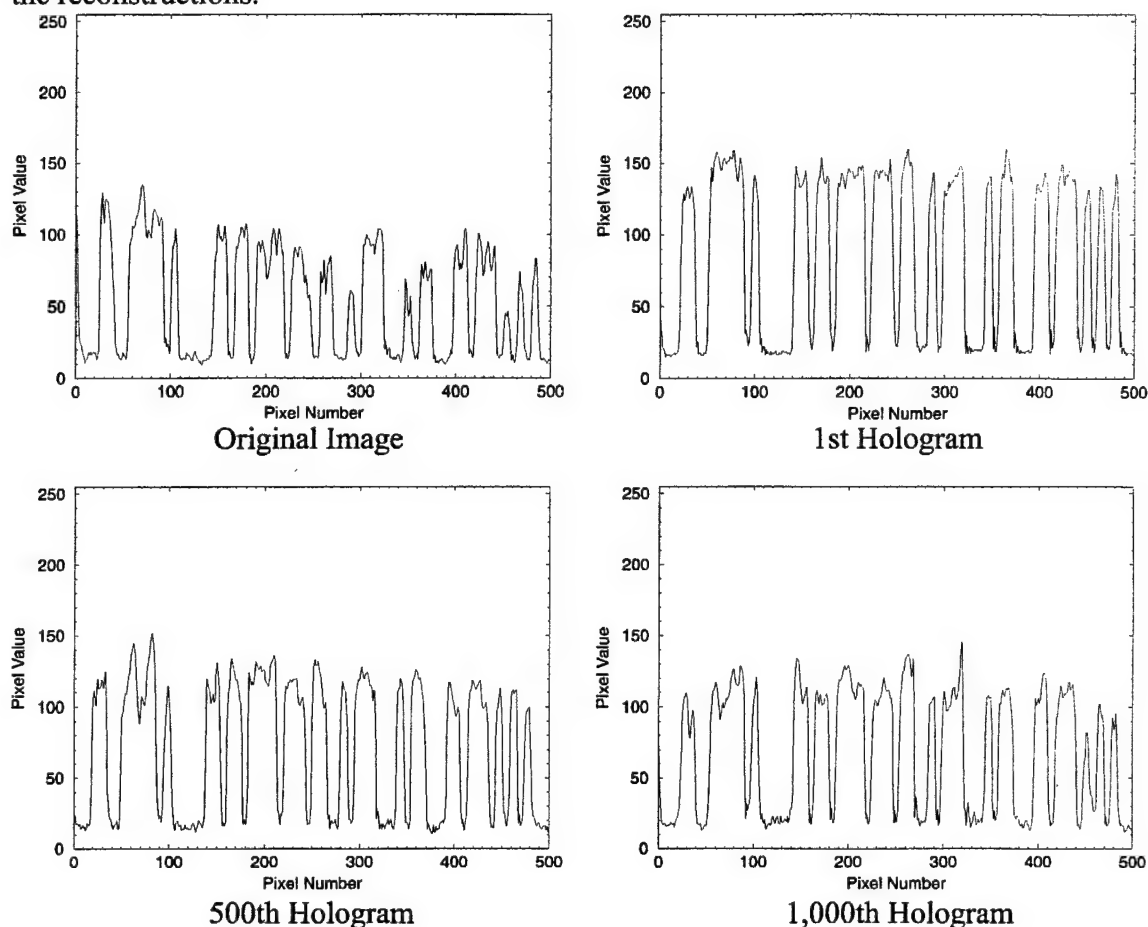Figure 2.34 : An image of the random data mask used to record 1,000 holograms and 3 of the reconstructions.

Original Image

1st Hologram

500th Hologram

1,000th Hologram

Figure 2.35 : The cross-section of the original and 3 reconstructed holograms for the 1,000 hologram experiment.

2.5 Discussions and Conclusions

In this chapter, several different recording materials for holography were discussed. However, for the application of mass distribution of pre-mastered information, only the DuPont photopolymer is well suited. Due to the limited thickness of DuPont's photopolymer, a new multiplexing method was invented to maximize its usefulness. Peristrophic multiplexing increased the number of holograms that can be stored in the

DuPont photopolymer by more than an order of magnitude. This shifts the limiting factor for the maximum number of holograms multiplexable in a given system from system design to the recording material. In order to fully realize the potential of the photopolymer, we characterized its recording behavior with recording curves. These recording curves show the dependence of grating formation on the exposure energy. A recording schedule to store equal strength holograms, using the photopolymers, can be derived from the recording curves. The recording schedule uses all the available dynamic range and takes into account the setup and "run-time" considerations. Using the schedule, we determined that the diffraction efficiency scales as $(2.2/M)^2$ and $(6.5/M)^2$ for DuPont's HRF-150-38 and HRF-150-100 photopolymers respectively, where M is the number of holograms stored. This predicts that 1,000 holograms can be multiplexed at a single location using the HRF-150-100 photopolymer for diffraction efficiency of $\sim 4 \times 10^{-5}$. This was verified when 1,000 holograms of a random bit pattern were recorded using the schedule process and the resulting average diffraction efficiency matched very well with the $M/\#$ prediction.

References

[1] Agfa-Gevaert 8E75 HD NAH glass plate films.

[2] A. Ashkin, G. D. Boyd, J. M. Dziedzic, R. G. Smith, A. A. Ballman, J. J. Levinstein, and K. Nassau. "Optically-induced refractive index homogeneities in LiNbO$_3$ and LiTaO$_3$." *Applied Physics Letters*, 9(1):72-74, 1966.

[3] F. S. Chen, J. T. LaMacchia, and D. B. Fraser. "Holographic storage in lithium niobate," *Applied Physics Letters,* 13(7):223-225, 1968.

[4] G. W. Burr, X. An, F. H. Mok, and D. Psaltis. "Large-scale rapid access holographic memory," *Optical Data Storage Meeting 1995*, SPIE Vol:2514-39.

[5] D. L. Staebler, W. J. Burke, W. Phillips, and J. J. Amodei. "Multiple storage and erasure of fixed holograms in Fe-Doped LiNbO$_3$," *Applied Physics Letters*, 26(4):182-184, 1975.

[6] F. Micheron and G. Bismuth. "Electrical control of fixation and erasure of holographic patterns in ferroelectric materials," *Applied Physics Letters*, 20(2):779-81, 1972.

[7] S. Li, F. Mok, and D. Psaltis. "Non-volatile holographic memory in photorefractive crystals using two wavelengths," *Applied Optics*, 19(3):210-212, 1994.

[8] Y. Qiao, D. Psaltis, and C. Gu. "Phase-locked sustainment of photorefractive holograms using phase conjugation," *Journal of Applied Physics*, 70(8):4646-4648, 1991.

[9] J. A. Jenny. "Holographic recording with photopolymers," *Journal of the Optical Society of America*, 60(9):1155-1161, 1970.

[10] D. A. Waldman, R. T. Ingwall, P. K. Dhal, M. G. Horner, E. S. Kolb, H. Y. S. Li, R. A. Minns, and H. G. Schild. "Cationic ring-opening photopolymerization methods for volume hologram recording," Unpublished Polaroid internal paper.

[11] K. Shutter and P. Gunter. "Photorefractive gratings in the organic crystal 2-cycloctylamino-5-nitropyridine doped with 7, 7, 8, 8-tetracyanoquinodimethane," *Journal of the Optical Society of America B*, 7(12):2278, 1990.

[12] B. Volodin, Sandalphon, K. Meerholz, B. Kippelen, N. V. Kukhtarev, and N. Peyghambarian. "Highly efficient photorefractive polymers for dynamic holography," *Optical Engineering*, 34(8):2213-2223, 1995.

[13] B. L. Booth. "Photopolymer material for holography," *Applied Optics*, 14(3):593-601, 1975.

[14] W. K. Smothers, T. J. Trout, A. M. Weber, and D. J. Mickish. "Hologram recording in DuPont's new photopolymer material," *2nd International Conference on Holographic Systems*, Bath, UK (1989).

[15] A. M. Weber and T. J. Trout. "Photopolymers for holography and other applications," *Proceedings of IUMRS-ICAM-93*, Tokyo, Japan (1993).

[16] K. Curtis and D. Psaltis. "Recording of multiple holograms in photopolymer films," *Applied Optics*, 31(35):7425-7428, 1992.

[17] F. H. Mok, M. C. Tackitt, and H. M. Stroll. "Storage of 500 high-resolution holograms in a LiNbO$_3$ crystal," *Optics Letters*, 16 (8):605-607, 1991.

[18] F. H. Mok. "Angle-multiplexed storage of 5000 holograms in lithium niobate," *Optics Letters*, 18(11):915-917, 1993.

[19] F. H. Mok, G. W. Burr, and D. Psaltis. "Angle and space multiplexed random access memory (HRAM)," *Optical Memory and Neural Networks*, 3(2):119-127, 1994.

[20] K. Curtis, A. Pu, and D. Psaltis. "Method for holographic storage using peristrophic multiplexing," *Optics Letters*, 19(13):993-994, 1994.

[21] H. Y. S. Li and D. Psaltis. "Three-dimensional holographic disks," *Applied Optics*, 33(17):3764-3774, 1994.

[22] K. Curtis. "Phase grating profiles in photopolymers," *Optics Communications*, 95(13):113-116, 1993.

[23] K. Curtis and D. Psaltis. "Characterization of the DuPont photopolymer films for 3-D holographic memories," *Applied Optics*, 33(23):5396-5399, 1994.

[24] U. Rhee, H. J. Caulfield, J. Shamir, C. S. Vikram, and M. M. Mirsalehi. "Characteristics of the DuPont photopolymer for angularly multiplexed page-oriented holographic memories," *Optical Engineering*, 32(8):1839-1847, 1993.

[25] G. Zhao and P. Mouroulis. "Diffusion model of hologram formation in dry photopolymer materials," *Journal of Modern Optics*, 41(10):1929-1939, 1994.

[26] D. Brady and D. Psaltis. "Control of volume holograms," *Journal of the Optical Society of America* **A**, 9(7):1167-1182, 1992.

[27] A. Papoulis. *Random Variables, and Stochastic Processes*, (McGraw-Hill, New York, 1991), Third Edition.

Chapter 3 : High Density Holographic Data Storage

3.1 Introduction

Holographic data storage offers several advantages when compared to conventional storage devices. The data stored and retrieved is organized as a two-dimensional page, consisting of a large number of pixels (bits). The read-out rate can be very high since an entire page of data is presented when a hologram is reconstructed. For example, accessing a page of data consisting of 1,000 × 1,000 pixels within 1 ms would give a data transfer rate of 1 gigabits per second. Furthermore, holographic data storage allows many pages of data to be stored at the same location with little cross-talk by using multiplexing techniques such as angle [1, 2], wavelength [3, 4], phase-code [5, 6], fractal [7, 8], peristrophic [9], and shift [10]. To a first-order approximation, the density achievable with holography is simply :

$$D_{3D} = M \times D_{2D} \tag{3.1}$$

where M is the number of holograms multiplexed at one location, D_{2D} is the density per page of hologram, and D_{3D} is the holographic density. For thin recording materials such as the photopolymer, it is convenient to think of the 3-dimensional holographic density as a surface density. This allows easy comparison with conventional data storage devices.

The surface density of current compact disc memories is approximately 1 bit/μm^2 and is limited primarily by the size of the illuminating spot. The next generation of optical compact discs that are scheduled to appear in the near future will have a surface density of ~6 bits/μm^2 [11]. Dual-layer and double-sided systems are expected to have an equivalent surface density in excess of 20 bits/μm^2. For holographic memories to be competitive, the surface density of holographic 3-D disks must be higher than the projected density of conventional media by a comfortable margin. This concern makes surface density the primary goal in the design of a holographic 3-D disk system. The storage density of holographic disks was analyzed in great detail by Li and Psaltis [12]. For the system they had in mind, it was determined that for an optimal recording medium thickness of 1.6 cm, a storage density of 110 bits/μm^2 is obtainable by angle multiplexing alone. This gives nearly an order of magnitude of margin over the projected surface density of the next compact disc technology.

Much in the thinking of holographic data storage system design has changed since Li and Psaltis' paper. New multiplexing methods such as peristrophic and shift have increased the storage density possible while using thinner recording materials. The old thinking of more holograms stored at a given location equals higher storage density has been proven to be only partially true. In Chapter 2.4.5, the experiment where 1,000 image plane holograms were stored in the 100 micron thick DuPont photopolymer was discussed. The size of each data pixel was 100 μm × 100 μm, giving a storage density of only ~.1 bit/μm^2 (1,000 holograms / (100 μm × 100 μm)). This is about an order of magnitude lower than the surface density of compact discs despite being the largest number of holograms ever stored in a 100 micron thick recording material. In Chapter

3.2 we will show an experiment where a surface density of 10 bits/μm^2 was achieved in the same 100 micron thick photopolymer by recording only 32 high bandwidth holograms.

The 10 bits/μm^2 experiment was to demonstrate that high density holograms could indeed be recorded in DuPont's photopolymer films with sufficient signal-to-noise ratio (SNR). In Chapter 3.3, we make our first attempt at constructing a functional holographic 3-D disk system using the DuPont photopolymer. The same 10 bits/μm^2 experiment was repeated except that the recording material was in the shape of a disk and shift multiplexing was used. Shift multiplexing is especially well suited for the holographic 3-D disk application. It allows holograms to be stored sequentially on the disk, in a spiral track much like the compact disc. The holograms partially overlap in the track and adjacent tracks also overlap to further increase the surface density. All this is done with a simple spherical reference beam and the disk motion provides the means of multiplexing. A very nice feature of using a spherical reference beam is the ability to diminish the effect of photopolymer shrinkage on the stored holograms. This is discussed in detail in Chapter 3.3.3.3.

A surface density of 10 bits/μm^2 is not enough to make holographic 3-D disks more attractive than conventional compact discs. To be truly impressive, a demonstrated surface density of 100 bits/μm^2 would be nice. However, this requires a thicker recording medium than the 100 micron thick photopolymer. Chapter 3.4 shows an experiment where a surface density of 100 bits/μm^2 was demonstrated with a 1 mm thick LiNbO$_3$, using nearly the same setup as in Chapter 3.3. This experiment shows that it is possible

to achieve a holographic surface density that is much higher than conventional data storage devices.

A lot of kinks remain to be ironed out before holographic 3-D disks can become a commercial product. Chapter 3.5 discusses some of the technical problems and possible solutions.

3.2 10 bits/μm^2 High Density Experiment Using Peristrophic and Angle Multiplexing

3.2.1 Introduction

Eq. 3.1 shows that the holographic storage density is equal to the number of holograms multiplexed at a given location times the page density of a single hologram. The same storage density can be achieved with many low page density holograms or a few high page density holograms. Given the desired storage density, the choice between many low density holograms or a few high density holograms is a complex issue. It depends on the system components available, the desired page size (number of pixels or area), the thickness and dynamic range of the recording material.

System components such as lenses and spatial-light modulators usually define the page density while the number of holograms is limited by the thickness and the dynamic range of the recording material. Since the maximum storage density is fixed for a given volume, there is also some cross relationship between page density and the number of holograms storable at a single location. The number of pixels in each hologram is usually picked to optimize the performance of the system. Depending on the application, it might

be desirable to have the memory space divided into large or small pages. For example, the Intel 8088 processors organized its 1 megabyte memory space into 64 kilobyte pages for easy access. This caused many headaches for programmers but was most efficient for the processor.

For the high density experiments in this chapter, we maximized the page density first and then tried to store as many holograms as possible in the recording material. This method uses the dynamic range of the recording material most efficiently since diffraction efficiency is inversely proportional to the number of holograms squared ($\eta \propto 1 / M^2$). The pixel size was picked to ease data readout and the number of pixels in each hologram was defined by the lens aperture. Using our high density setup, we were able to achieve a surface density that is many times higher than any conventional data storage devices.

3.2.2 Experimental Setup

Figure 3.1 shows a schematic diagram of the high density setup using peristrophic and angle multiplexing. A photographic glass plate of a random binary bit pattern was used as the input SLM. The center-to-center spacing of the pixels was 45 microns (in both directions), and the fill factor was 100%. A pair of Nikon f/1.4, 3.9-cm-aperture camera lenses were used in a 4-f geometry to image the SLM to the detector array. The glass data mask plate was pressed up against the Nikon lens to ensure that all the pixels within the lens aperture were captured and imaged to the detector plane. A total of 590,000 pixels fit in the apertures of the two Nikon lenses and a sharp image of the entire field was obtained at the detector plane. The holograms were recorded by a planewave reference beam with the recording material placed slightly past the Fourier transform

plane. At that position, the diameter of the signal beam on the photopolymer was 1.5 mm and its spatial uniformity was much better than at the exact Fourier plane. The aperture of the reference planewave on the recording material was controlled by an iris. To avoid diffraction effects, the iris was imaged to the photopolymer by a 4-f system with the iris at the first image plane and the photopolymer at the other image plane. The intensity of the reference planewave was 250 μW/cm^2 and the intensity of the signal planewave before the data mask was 1.8 μW/cm^2. The photopolymer was tilted by 5° out of the x-y plane to reduce back reflections. Peristrophic multiplexing was achieved by rotating the photopolymer around its own surface normal. Angle multiplexing was combined with peristrophic multiplexing to increase the number of pages stored at a single location. For simplicity, angle multiplexing was achieved by rotating the recording material around the x-axis with a second rotational stage. The signal and reference beams made an angle of 25° and 35° (respectively, measured from outside the photopolymer) with respect to the surface normal of the photopolymer initially. The rotation of the recording material around the x-axis during angle multiplexing changes these angles but the angle between the signal and reference beams remains constant at 60°. At the detector plane, a simple video rate CCD camera was used. Its output was linked to a computer via a framegrabber board.

Figure 3.2 shows a picture of the high density setup using peristrophic and angle multiplexing. The dimensions of the setup, measured from the front of the data mask to the end of the CCD was approximately 28 cm with a width of 12 cm. A photographically produced data mask was used in the setup instead of a liquid-crystal type SLM for several reasons. (1) The 100% fill factor of the photographically produced data mask eliminated

the repetitive higher orders. This simplified the setup since no filtering was required. (2)

To ensure that the highest possible density is achieved in a page of hologram, the entire

aperture of the Nikon lens must be filled with data. The liquid-crystal SLMs we had in

the labs were too small to cover the entire 3.9 cm Nikon lens aperture. Therefore, a large

but fixed data mask was used in the setup. Since this is a first-order experiment to

demonstrate the high storage density capability of holographic devices, we feel that the

inability to change the data pixels from page to page is but a small kink in the experiment.
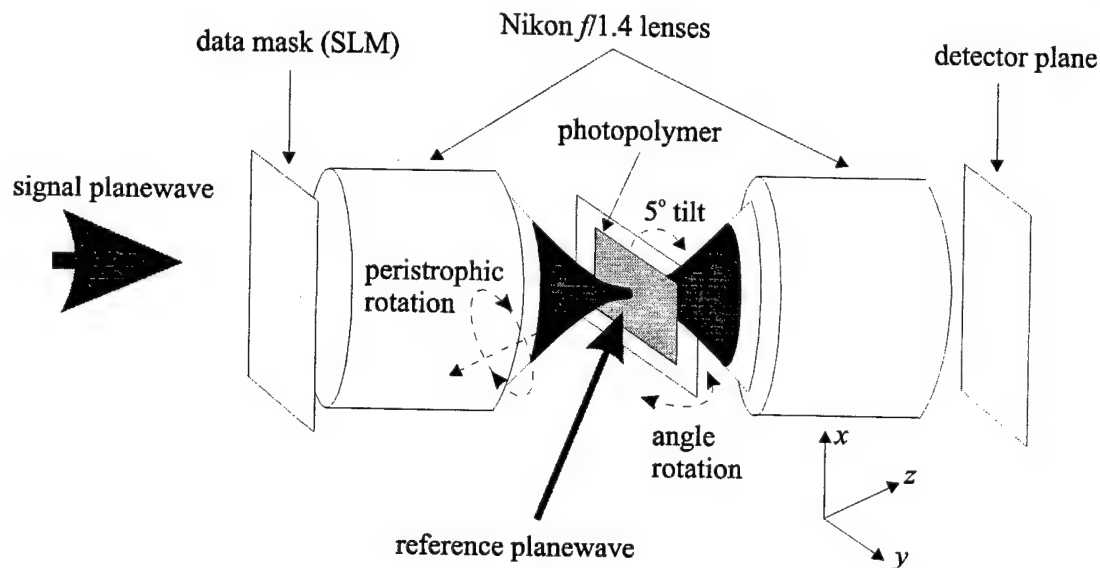

Figure 3.1 : Schematic diagram of the high density setup using peristrophic and angle
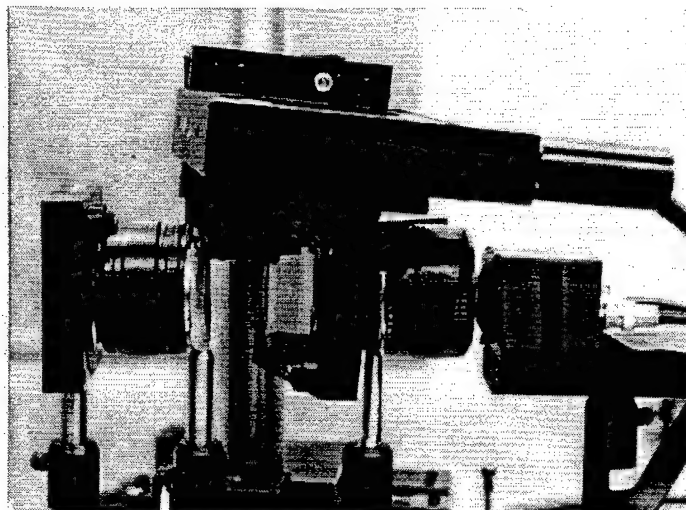multiplexing.

Figure 3.2 : A picture of the high density setup using peristrophic and angle multiplexing.

Nikon camera lenses were used in the setup because they are relatively inexpensive for a multi-element lens and they are very well corrected across the entire aperture. For imaging applications where a low $f/\#$ system is required, significant image distortion usually results at the edges of the lenses. Camera lenses are designed to faithfully reproduce the input scenes on films without much distortion. This is achieved by using multiple lens elements in a single package. For example, the Nikon $f/1.4$ lens (Product Number 1904) has 7 elements in 6 different groups to correct for distortions. Camera lenses are produced for the mass market, therefore they are relatively inexpensive (~\$400 for a Nikon $f/1.4$ lens) when compared to custom designed lenses with the same specification (~\$30,000 for the design and production of the first lens). No significant image distortion was encountered anywhere within the apertures of the Nikon lenses. .

Notice in Figure 3.1, the 4-f system constructed with the two Nikon $f/1.4$ camera lenses is actually backwards. To construct a true 4-f system in the way these lenses were designed, the side of the lens that normally faces the photographic film should face the data mask and the CCD array. There is a separate ring that allows two Nikon camera

lenses to be screwed together to form a perfect 4-f system. The ring screws the sides of the lenses that normally face the scene together. The problem with that setup is then the Fourier plane becomes inaccessible since the focal plane of the side that normally faces the scene is only a few millimeters from the lens casing. On the other hand, the focal plane of the side that faces the photographic film is about 4 cm from the lens casing. By building the 4-f system backwards, we can get almost 8 cm of working space between the lenses, enough room to place the recording material and the rotational stages near the Fourier plane. This backward 4-f system undoubtedly has some aberrations. However, it was not noticeable with the large 45 μm $\times$ 45 μm pixels in our data mask.

We chose near Fourier plane recording instead of the image plane because it is easier to achieve high page density holograms in the Fourier plane. Figure 3.3 shows a typical image plane setup using two 4-f imaging systems. The data to be stored is first imaged to the photopolymer by the first 4-f system. The interference pattern created between the image of the data mask and a planewave reference beam is stored as a hologram. During readout, the reconstructed hologram is imaged to the detector array with the second 4-f system.
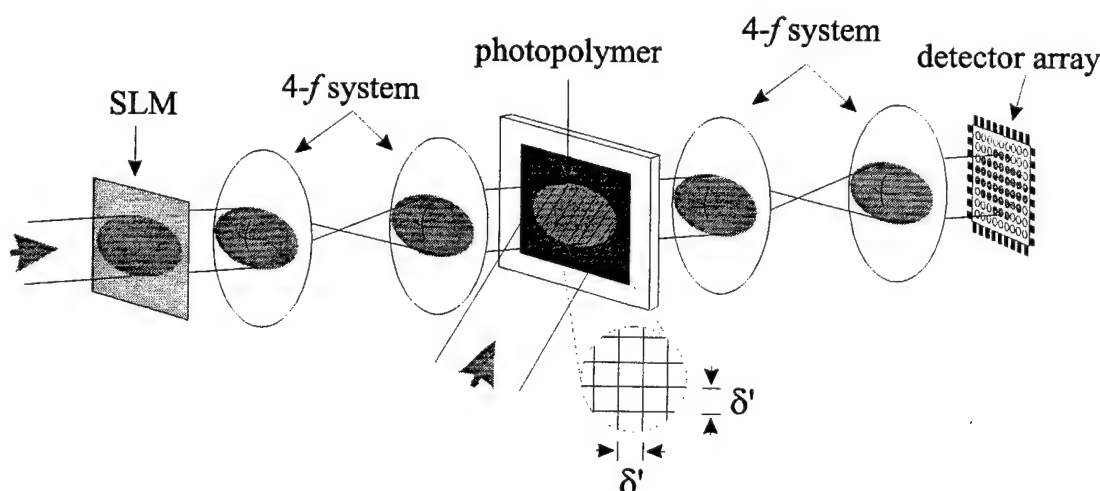
Figure 3.3 : Typical image plane recording setup.

In the image plane, the density per page is given by the inverse of the pixel area (density $\simeq 1/\delta'^2$). Therefore, in order to achieve high density per page, the pixel size must be very small. However, the pixel width of most SLMs and detector arrays are usually from slightly less than 10 μm to tens of microns. This gives a very low density value unless the first 4-f system reduces the pixel size and the second 4-f system re-expands the pixels. Lens systems that can reduce input pixels to the micron range are very expensive. Some lithographic reduction lenses manufactured for the VLSI industry can be applied to this application. For example, the Nikon Ultra-Micro-Nikkor 28 mm f/1.7g ultra high resolution lens was designed for printing minute image patterns on plates coated with photo-resist under g-line monochromatic lighting (435.8 nm wavelength). Its standard reduction is 1/10 and this lens guarantees a resolving power of more than 800 lines per mm over the 8 mm image diameter (minimum pixel width of 1.25 μm). Using two of these lenses to image the data from the input SLM to the detector array would give a maximum page density of .64 bits/μm^2 at the reduced image plane between the two lenses. This density value is very good considering that the same resolution is maintained over a large field (unlike compact discs where the objective lens only needs to produce a single spot at the optical axis of the lens). However, these lenses are very specialized, expensive, and sometimes even unobtainable. Newer lens systems that can resolve sub-micron pixels are currently exposing the next generation VLSI chips with feature sizes less than .2 microns. However, these lithographic systems are sold with million dollar price tags and the lenses cannot be brought separately.

Another draw back of these high resolution reduction lens systems is that their overall working distance is large. For the Nikon Ultra-Micro-Nikkor 28 mm f/1.7g lens, the distance from the input image to the reduced image is 31.5 cm. Using two of these lenses to reduce and re-expand the input image would require a system with a length of at least 63 cm. Furthermore the distance from the lens casing to the reduced image plane is only 1.1 cm. This gives a distance of only 2.2 cm between the two lenses to place the rotational stages and the recording material.

For Fourier plane recording, the criteria for high density holograms is completely different. Figure 3.4 shows a simple Fourier plane holography setup using one 4-f system. At the Fourier plane of the 4-f system, the width of the signal beam is given by $2\lambda f/\delta$, where λ is the wavelength, f is the focal length of the lens, and δ is the pixel size of the SLM. For large pixels, the spot at the Fourier plane will be small and vice versa. The density per page at the Fourier plane is equal to the number of pixels captured by the apertures of the 4-f system over the spot size of the hologram. For a square SLM with N_p by N_p number of pixels that fit within the lens aperture, the page density is given by :

$$\text{page density at the Fourier plane} = \frac{N_p^{\,2}\delta^2}{(2\lambda f)^2} \tag{3.2}$$

where $N_p\delta$, the number of pixels per dimension times the pixel size, can be thought of as the maximum aperture of the lens system. Then $N_p\delta/f$ is like the inverse f/# of the lens system. By plugging $f/\# = f/N_p\delta$ into Eq. 3.2 and multiply the whole thing by a circular lens aperture correction factor, we get :

$$\text{page density at the Fourier plane} = \frac{1}{(2\lambda f/\#)^2}\frac{\pi}{4} \tag{3.3}$$

where *f*/# is the *f*-number of the 4-*f* system. Notice in Eq. 3.3, the page density at the Fourier plane is independent of the SLM pixel size. This means that larger pixels can be used and still get the maximum possible page density given the *f*/# of the 4-*f* system. Hence the pixel size can be picked to match conveniently with the pixel size of the detector array without expansion or reduction. Furthermore, when compared with the image plane setup, the Fourier plane system is much more compact and fewer lenses are required.

The only requirement on the lens system for the Fourier plane setup is that the lenses be well corrected for distortion. If the pixels at the edges of the lenses get warped or distorted in some way, then the actual *f*/# of the 4-*f* system is higher because those pixels cannot be used. This decreases the density possible per page. Luckily, camera lenses are very well corrected across the entire field of view and some have low *f*/#'s. For the Nikon *f*/1.4 camera lenses used in this experiment, the maximum possible density per page at the Fourier plane, using a wavelength of 532 nm would be .35 bits/μm^2. Even higher density can be achieved with lower *f*/# camera lenses. We also tried to use Nikon *f*/1.2 camera lenses (Product Number 1435). However for these lenses, the focal plane of the side that faces the scene is inside the lens casing. This means that when used as a 4-*f* system, the image plane is inaccessible. So far we have not found any other lower *f*/# camera lenses with both focal planes outside the lens casing.
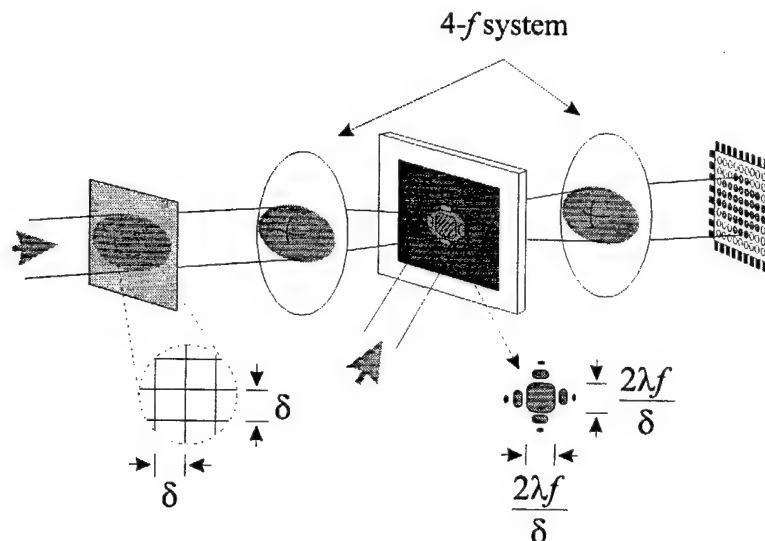
4-*f* system

Figure 3.4 : A simple Fourier plane holography setup.

The ability to pick convenient pixel sizes (and still maintain the same page density) and readily available camera lenses that can be used to construct a well corrected 4-*f* system are the two major reasons why the Fourier plane geometry is so attractive. In the limit of the resolvable spot size given the *f*/# of the lens system ($\delta = 2\lambda f/\#$), the density achievable in the image plane is the same as in the Fourier plane. However, Fourier plane holography offers better immunity to dusts, scratches, and surface defects since the energy of each pixel is spread out over a larger area.

On the other hand, there is a big problem with recording holograms at the exact Fourier plane. At the Fourier plane, the signal beam has a very strong DC component. The DC spike at the Fourier plane is from the non-spatially varying part of the input image. If a hologram was recorded at the exact Fourier plane, the strong DC quickly exhausts the dynamic range of the photopolymer at the DC spot. This causes subsequent holograms that are recorded at the same location to become edge enhance, since only the higher frequency components are stored. One method to get around this problem is to use

a random phase diffuser [13-15]. A phase diffuser placed right before or after the SLM modulates the phase of the input image. The non-spatially varying part of the input image gets modulated by a random phase, which causes the DC to broaden in area. If the resolution of the random phase diffuser is better than the SLM, then the Fourier transform spot would get bigger since the effective δ has gotten smaller. This is undesirable since the density per page would decrease due to the increased hologram area. However, by making the random phase diffuser pixel matched with the SLM pixel, the same page density can be obtained with a more uniform Fourier transform.

In practice, it is often difficult or costly to obtain a good random phase diffuser. We tried to make binary $(0, \pi)$ random phase diffusers in-house for the high density experiments. However, we were unable to achieve the desired effects due to poor etch uniformity and rough pixel alignments. For the experiments in this chapter, instead of using a phase diffuser, we move the recording material slightly past the exact Fourier plane so that the DC component could defocus and become more uniform. Figure 3.5 shows a simple graphical representation of how the rays converge at the Fourier plane.
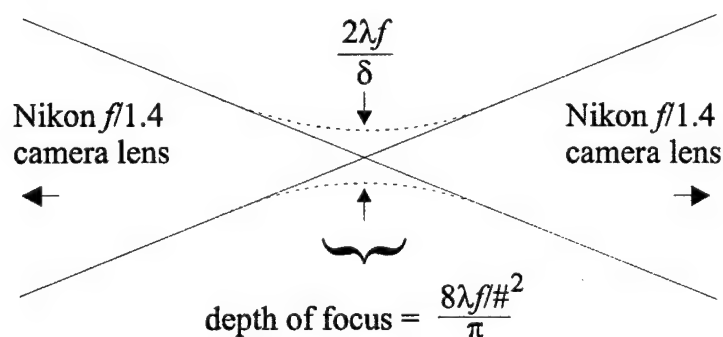

Figure 3.5 : Graphical representation of how the rays converge at the Fourier Plane.

The Nikon $f/1.4$ camera lenses produce a cone of rays nearly $40°$ in spread. The DC component focuses down to a waist size of a few microns while the higher frequency

components produce a zero[th] order waist of $2\lambda f/\delta$. Approximately 90% of the energy is in the zero[th] order with the remaining 10% in higher orders. The depth of focus of the DC spot (when the waist becomes $\sqrt{2}$ times larger) is given by the expression shown in Figure 3.5. For the small $f/\#$ Nikon lenses used, the DC spot defocuses very quickly and catches up with the higher frequency components with only a small displacement off of the Fourier plane. At the exact Fourier plane, the waist size of the zero[th] order is 1.3 mm ($\lambda = 532$ nm, $f = 5.5$ cm, and $\delta = 45$ μm). For the high density experiments, we slightly displaced the recording material past the exact Fourier plane and the diameter of the signal spot on the photopolymer was 1.5 mm when sufficient DC spreading was achieved. In measuring this spot size, the very weak higher frequency components in the signal were excluded. These low intensity components do not record well due to their poor modulation index. Therefore, they could be filtered during recording or just not illuminated during readout to keep the effective area of the hologram small. This low-pass filtering process causes the edges of the reconstructed pixels to be not quite as sharp as the original data mask. However, the smaller signal spot size does increase the page density. The density per page at the slightly past Fourier plane position is 590,000 pixels / $(0.75 \times 0.75 \text{ mm}^2 \times \pi) = 0.33$ bits/μm². Not as good as the maximum possible density of 0.35 bits/μm² at the exact Fourier plane but pretty close.

3.2.3 Experimental Results

The goal of this experiment is to achieve an equivalent surface density of 10 bits/μm² in DuPont's HRF-150-100 (100 microns thick) photopolymer using angle and peristrophic multiplexing. With a page density of .33 bits/μm², a minimum of 30

holograms must be multiplexed in the same area to get a surface density of 10 bits/μm^2. This might sound like a simple proposition since 1,000 holograms were stored in the same recording material, using the same multiplexing methods in Chapter 2.4.5. However, to store 30 holograms in this experiment will prove more difficult due to the much higher bandwidth of the signal beam.

For the 1,000 hologram experiment shown in Chapter 2.4.5, 100 peristrophic holograms were multiplexed at each of the 10 angle multiplexing positions (position here refers to the angular position of the rotational stage, all the holograms were stored at one location). The photopolymer was rotated in-plane by 1.8° for peristrophic multiplexing and angular positions were separated by 1.5°. For our high density experiment, the aperture of the Nikon f/1.4 camera lens is 3.9 cm and the focal length is approximately 5.5 cm. The signal and reference beams made an angle of 25° and 35° (respectively, measured from outside the photopolymer) with respect to the surface normal of the photopolymer. The theoretical Fourier plane peristrophic selectivity as computed from Eq. 2.5 (where $N_p\delta$ = 3.9 cm) is a minimum of 41 degrees in rotation to shift the reconstruction completely out of the aperture of the Nikon lens. Experimentally, we found that a 45° in rotation was sufficient to completely Bragg mismatch and shift the undesired hologram off of the Nikon lens aperture. Therefore, up to four peristrophic holograms can be multiplexed from 0 to π with this configuration. The small difference between theory and experiment is probably due to the slight displacement of the recording material off of the exact Fourier plane.

Due to the much higher bandwidth signal beam in the high density experiment, instead of being able to store 100 peristrophic holograms from 0 to π (as in the 1,000

hologram experiment), only four can be stored without cross-talk. The reconstructed hologram in the high density setup is much harder to get rid of since it is so large (3.9 cm in diameter) and the focal length of the Nikon camera lenses are so short (5.5 cm). We could have used longer focal length lenses to decrease the amount of rotation required to shift away the undesired hologram. However, the density per page would decrease if the aperture of the new camera lens was not bigger to capture more pixels (since the area of the hologram would increase with the longer focal length). On the other hand, by increasing the aperture of the new lenses, the undesired hologram would again become harder to get rid of. We can derive the surface density achievable with peristrophic multiplexing based on the $f/\#$ of the imaging system.

Fourier plane peristrophic selectivity (Eq. 2.5) can be rewritten as :

$$d\psi = \frac{1}{f/\#\left(\sin\theta_r + \sin\theta_s\right)} \tag{3.4}$$

by replacing $f/N_p\delta$ by the $f/\#$ of the imaging system. The number of holograms that can be peristrophically multiplexed from 0 to π in rotation is given by :

$$\text{number of peristrophic holograms} = \pi f/\#\left(\sin\theta_r + \sin\theta_s\right) \tag{3.5}$$

Eq. 3.5 can be combined with Eq. 3.3 to get the surface density achievable with peristrophic multiplexing.

$$\text{surface density with peristrophic multiplexing} = \frac{\pi^2}{4}\frac{\sin\theta_r + \sin\theta_s}{(2\lambda)^2 f/\#} \tag{3.6}$$

From Eq. 3.5 we can see that if a lot of peristrophic holograms is what we want, then a large $f/\#$ lens system should be used. However, the surface density achievable at the Fourier plane is inversely proportional to the square of the $f/\#$. Therefore, in order to

get high surface density, the lowest $f/\#$ imaging system should be used (Eq. 3.6). This reduces the number of peristrophic holograms but increases the overall surface density.

The maximum number of holograms that can be angularly multiplexed at each location is given by :

$$\text{number of angle holograms} = \frac{\Theta}{\Delta\theta\,'} = \frac{\Theta L \sin(\theta_r\,'+\theta_s\,')}{\lambda\,'\cos\theta_s\,'} \tag{3.7}$$

where Θ is the range of angle that can be used to store angle multiplexed holograms, $\Delta\theta\,'$ is the angular selectivity inside the recording material (first null), L is the thickness of the hologram, $\lambda\,'$ is the wavelength of the laser inside the recording material, $\theta_r\,'$ ($\theta_s\,'$) is the angle between the reference (signal) beam and the photopolymer's surface normal measured from inside the material. For our experimental setup, $L = 100$ μm, the index of refraction of the photopolymer = 1.525, $\lambda\,' = 349$ nm, $\theta_r\,' = 22°$, and $\theta_s\,' = 16°$, which gives $\Delta\theta\,'$ a value of $0.31°$ to the first null of the sinc function inside the material. However, Eq. 3.7 is valid only for planewave holograms. The signal beam spread inside the recording material was nearly $25°$ because of the large numerical aperture of the $f/1.4$ Nikon lenses. The actual angular selectivity curve from one hologram stored using the high density setup is shown in Figure 3.6.
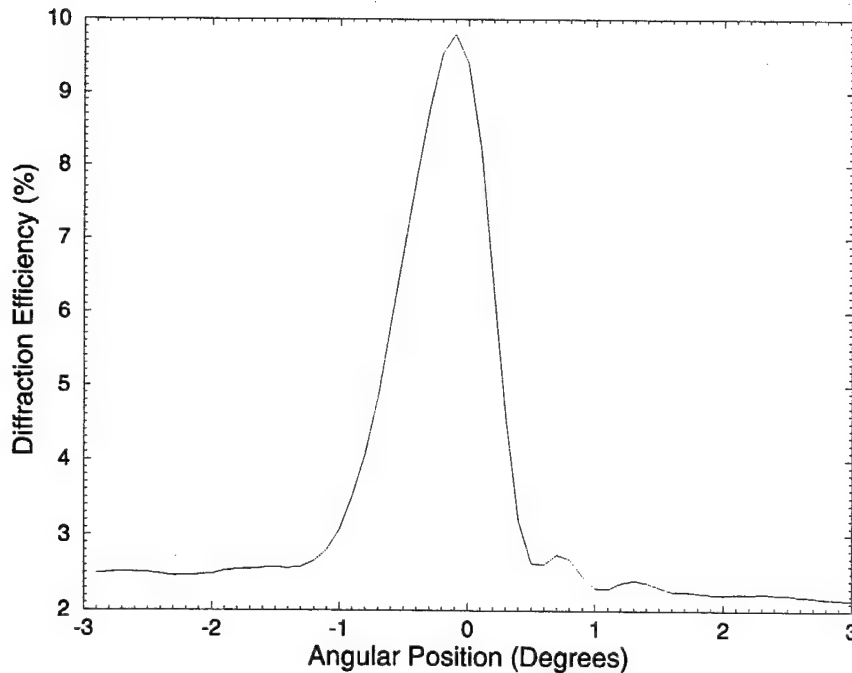
Figure 3.6 : The angular selectivity of one hologram recorded with $\theta_r' = 22^\circ$ and $\theta_s' = 16^\circ$.

The diffraction efficiency of the hologram was measured by placing a lens with sufficient aperture to cover the entire reconstruction at the image plane of the hologram. The lens collected the entire hologram and a silicon detector was placed near the focal plane of the lens to measure the diffracted power as the recording material was rotated to different angular positions. The hologram was stored with $\theta_r' = 22^\circ$ and $\theta_s' = 16^\circ$ and the '0' angular position in Figure 3.6 represents the angle of the rotational stage at which the hologram was stored.

Notice that the selectivity curve in Figure 3.6 is asymmetric and doesn't have the maximum diffraction efficiency at '0' angular position. This behavior does not match the simple angular selectivity theory of a sinc() function with regular null spacing. As stated in Chapter 2, DuPont's photopolymer shrinks slightly when exposed to light. This shrinkage is part of the chemical process that creates the index grating and causes the Bragg condition of the stored hologram(s) to change slightly. The shrinkage effect is

most apparent for high bandwidth holograms recorded in an asymmetric geometry.
Figure 3.7 shows the κ-sphere representation of a high bandwidth hologram in an
asymmetric geometry.

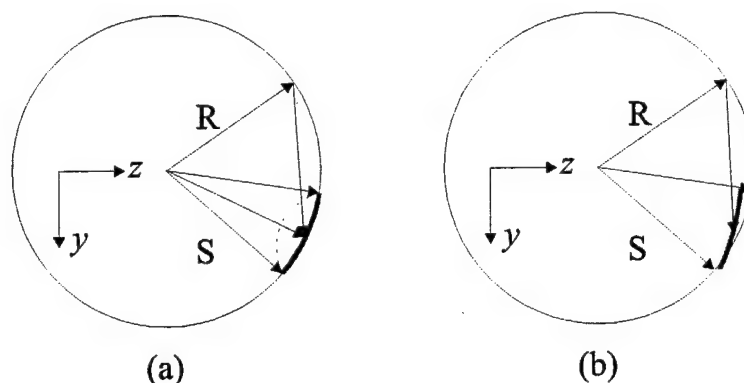Figure 3.7 : κ-sphere representation of a high bandwidth hologram in asymmetric
geometry. (a) The grating cone before shrinkage. (b) The grating cone after shrinkage.

Figure 3.7 (a) represents the κ-sphere diagram of a high bandwidth hologram
before shrinkage has occurred. The planewave reference beam is represented as a single
vector while the high bandwidth signal beam is represented as a spectrum of vectors.
There exists a grating vector from the tip of the reference beam vector to every signal
beam vector in the signal cone. One such grating vector is drawn in Figure 3.7 for
illustration and the grating cone on the κ-sphere in the y-z plane is highlighted with a
thick, curved line. Shrinkage of the photopolymer occurs mostly in the z direction since
the material is anchored on a glass substrate in the x-y plane. Photopolymer shrinkage
compresses the grating cone in the z direction and lifts it off of the κ-sphere as shown in
Figure 3.7 (b). The only part of the grating cone that remains on the κ-sphere is where
the signal beam has the same angle as the reference beam, with respect to the surface
normal of the photopolymer. For that special case, the grating vector has no z-
components and therefore shrinkage has no effect on its Bragg condition.

We can re-Bragg-match portions of the grating cone by changing the reference beam angle slightly or by rotating the photopolymer around the x-axis. Figure 3.8 shows how this is done in the κ-sphere. In Figure 3.8 (a), by rotating the photopolymer counter-clockwise around the x-axis, the reference beam angle would appear smaller and this pushes the grating cone away from the κ-sphere. With sufficient counter-clockwise rotation, the upper portion of the grating cone will get pushed back on the κ-sphere. In Figure 3.8 (b), by rotating the photopolymer clockwise around the x-axis, the reference beam angle would appear larger, and this pulls the lower portion of the grating cone back on the κ-sphere. Notice in either case, the portion of the hologram that is Bragg-matched diffract at a slightly different angle than the original image. Furthermore, only a portion of the hologram can be Bragg-matched at a time using this method. For the asymmetric recording geometry shown in Figure 3.8, the upper portion of the grating cone requires a larger rotation to Bragg-match than the lower portion. This causes the diffracted power to dissipate slower when rotated in the counter-clockwise direction. Also, since a larger portion of the grating cone requires a counter-clockwise rotation to Bragg-match, the maximum diffracted power angular position changes from no rotation to favor a slight counter-clockwise rotation. These effects are depicted in the angular selectivity curve of Figure 3.6. Negative angular positions refer to counter-clockwise rotation while positive angular positions refer to clockwise rotation of the photopolymer around the x-axis.
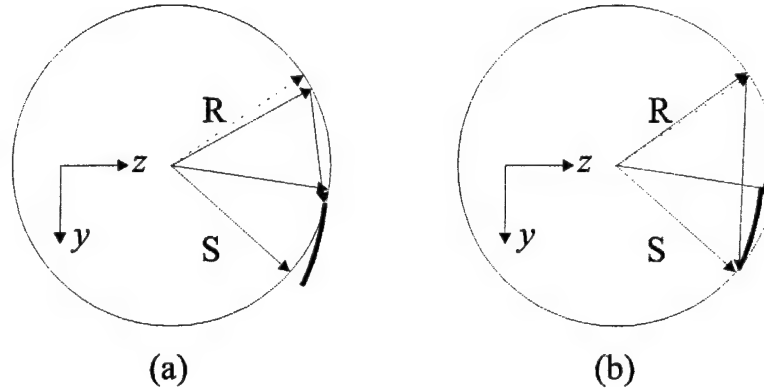
Figure 3.8 : κ-sphere representation of the Bragg-match condition. (a) To Bragg-match the upper portion of the grating cone. (b) To Bragg-match the lower portion of the grating cone.

We can figure out how much the photopolymer has shrunk by using the amount of rotation required to re-Bragg-match the different portions. Experimentally for our setup, the lower portion required almost no rotation to re-Bragg-match while the middle and the upper portions required a counter-clockwise rotation of 0.3 and 0.7 degrees, respectively. Assuming the modification to the reference beam angle is small when the photopolymer is rotated to re-Bragg-match different portions and the total index of refraction change is negligible, we can compute the shrinkage factor, γ, as :

$$\gamma = 1 - \frac{\cos\theta_s' - \cos\theta_r'}{\cos\theta_s' - \cos\left(\theta_r' + \Delta\theta'\right)} \tag{3.8}$$

where θ_r' is the reference beam angle, θ_s' is the signal beam angle for that portion of the hologram, $\Delta\theta'$ is the change in the reference beam angle due to the rotation (all angles measured from inside the recording material). By plugging in the numbers corresponding to the upper portion of the hologram ($\theta_r' = 22°$, $\theta_s' = 3.3°$, $\Delta\theta' = 0.4°$) we get a shrinkage factor of $\gamma = 0.04$. This means that the photopolymer shrunk by about 4% for that hologram.

It is hard to tell the null spacing from the angular selectivity curve of Figure 3.6 since the spacing is different for clockwise and counter-clockwise rotations. A better indication of where the next angle hologram can be multiplexed is the amount of noise introduced by adjacent angle holograms. This can be obtained by storing three angle multiplexed holograms and then measure the amount of noise experienced by the middle hologram. The SNR (using Eq. 2.11) measured from three angularly multiplexed holograms with an angular separation of $5°$ was 4 (a rotation of $5°$ around the x-axis). The quality of the reconstructions were good and no bit errors were observed. By decreasing the angular separation to $2°$, the SNR of the middle hologram dropped to 3.2 and the distinction between the 'on' and 'off' pixels were not as clear. No bit-errors were observed in the reconstruction of the middle hologram but errors were likely if the angular separation was decreased further. For comparison, the SNR measured for a single hologram (no multiplexing) was 4.5. Figure 3.9 shows the center portion of the reconstructed hologram for the three cases.

Figure 3.9 shows that visually, the quality of the hologram decreased as the angular separation was decreased from $5°$ to $2°$. In order to avoid bit-errors with this setup, the angular separation must be greater than $2°$. On the other hand, we would also like to store as many angle holograms as possible to boost the overall surface density. As stated before, with a page density of .33 bits/μm^2, we need at least 30 holograms to achieve a surface density of 10 bits/μm^2. We have already determined that up to 4 peristrophic holograms can be stored from 0 to π without cross-talk. In order to get at least 30 holograms, we would need to store 8 angle holograms at each of the 4 peristrophic positions. Our setup has a geometrically limited Θ (Eq. 3.7) of slightly less

than 20°. Therefore we picked an angular separation of 2.5° per hologram and hoped the SNR would be good enough to avoid bit-errors.
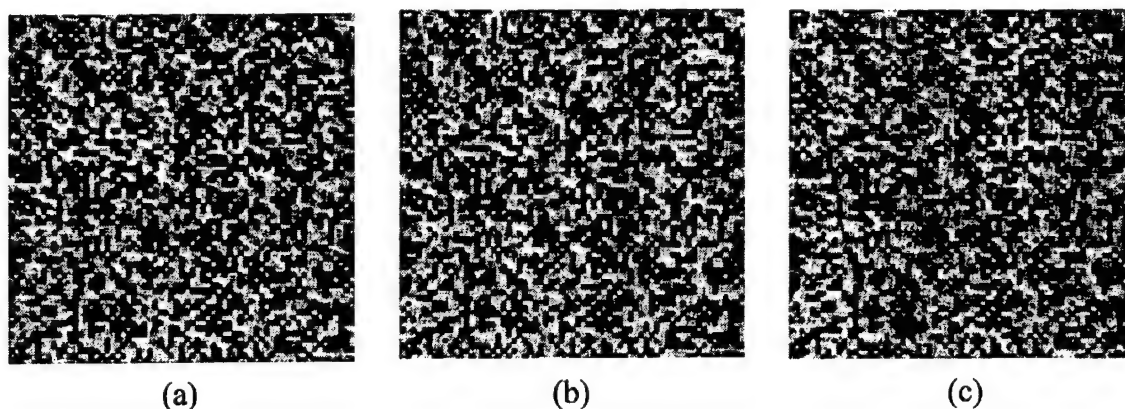

Figure 3.9 : The center portion of the middle hologram for three different cases. (a) One hologram. (b) Three angle multiplexed holograms separated by 5° each. (b) Three angle multiplexed holograms separated by 2° each.

The method in which we determine the amount of noise in a hologram is shown in Figure 3.10. The reconstructed hologram is viewed by a CCD detector array placed at the image plane. The active area of our CCD array (.65 mm × .48 mm) is much smaller than the aperture of the reconstructed hologram. Therefore, in order to get a fair representation of the quality of a hologram, the CCD array was physically moved to capture an image from the left, center, and right sides of the reconstructed hologram. The reported SNR of a hologram is usually the average SNR of the three images. From the approximately 140 × 100 data mask pixels captured within the CCD array aperture, a smaller image of 300 × 300 CCD pixels (about 65 × 65 data mask pixels) is sent to a computer program for analysis. Because the pixel pitch of the data mask is 45 μm and the pixel pitch of the CCD array is approximately 10 μm, each data mask pixel is over-sampled during readout. In order to get the true intensity values of the data mask pixels, a grid with the correct spacing is overlaid on the image and the intensity values of the CCD pixels within each

grid box is integrated to produce the data mask pixel value. The computer program first expands the image by 2X to generate an image with 600×600 CCD pixels so that a grid can be more easily placed. This also allows the user to specify a margin in which the edges of the grid box will not be integrated. In the un-expanded image, each data mask pixel is sampled by 20.5 CCD pixels on the average, and the program produces an estimate for the intensity of an data mask pixel by integrating only 9.5 CCD pixel values (on the average, at the center of the grid). Therefore, on the average, 11 CCD pixels are discarded as edge pixels. The computer program has *a priori* knowledge of which data pixels are supposed to be 'on' and 'off'. This allows the program to compute the SNR using Eq. 2.11, generate a histogram of the data mask pixels, and determine the bit-error-rate (BER) given a threshold value. The program can also 'auto-jiggle' the overlaid grid by a few CCD pixels to obtain the best SNR value possible.

After determining the necessary angular separation, sets of eight angularly multiplexed holograms ($2.5°$ separation) were recorded at each of the four peristrophic positions ($45°$ in rotation), for a total of 32 holograms stored. The overall surface density was 32 holograms $\times$ 0.33 bits/μm^2 or 10.6 bits/μm^2. The exposure time per hologram varied from 2 to 4 seconds and the exposure schedule was obtained iteratively through trail-and-error. We did not use the method discussed in Chapter 2.4.3 to compute the exposure schedule since only 32 holograms were required in this experiment. Figure 3.11 shows the diffracted powers of the 32 holograms. The average diffraction efficiency of the stored holograms was approximately 3.5×10^{-3}.

From C C D A rray

C o m p u ter

H isto g ram

S o ftw are

Combined Histogram of the 32 Reconstructions
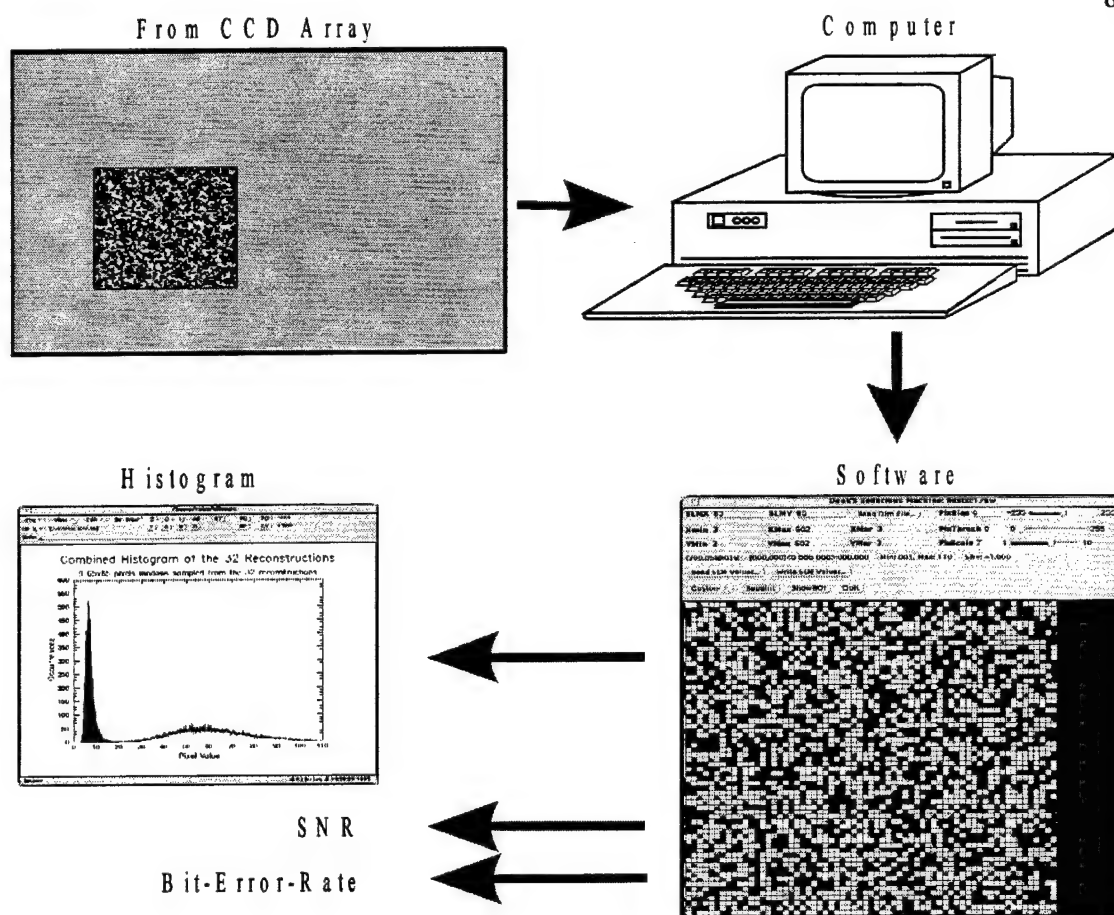
S N R

B it-E rro r-R ate

Figure 3.10 : Block diagram of how the stored data is read out and analyzed.

The left, center, and right sides of the data mask imaged through the system without any holograms can be compared with the reconstructions of some of the stored holograms in Figure 3.12. Visually, the quality of the stored holograms were fairly good and no bit-errors were found in the sampled holograms.
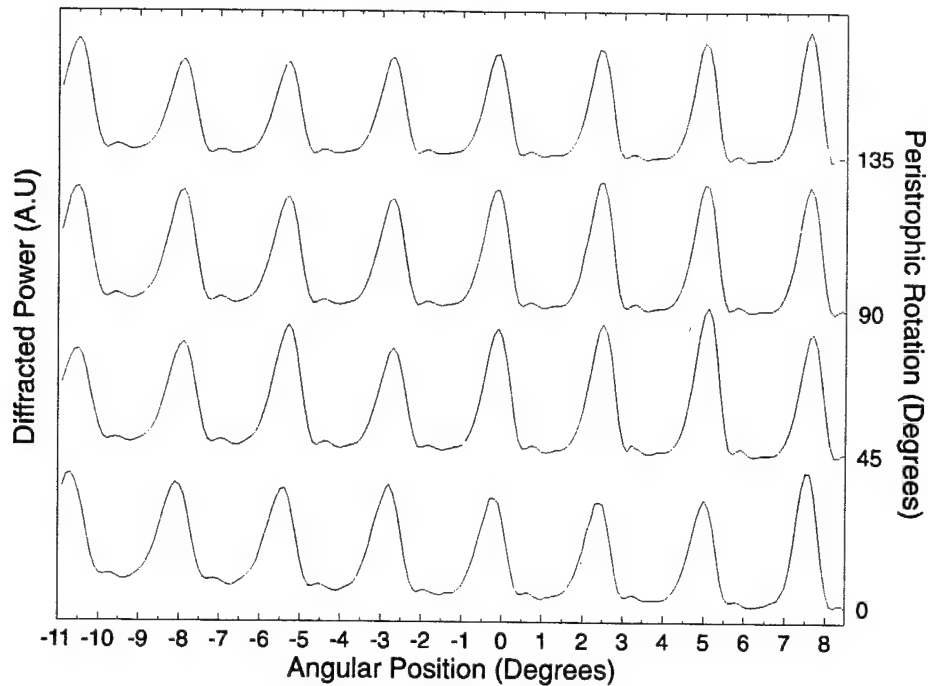
Figure 3.11 : Diffracted powers of the 8 angle multiplexed holograms at each of the 4 peristrophic positions.

The most challenging aspect of the high-density experiment proved to be the suppression of noise in the system in order to achieve acceptable levels of SNR and probability of error. The results shown in Figure 3.11 and 3.12 were achieved only after many iterations of fine tuning the setup. The way we achieved high density was by increasing the angular bandwidth of the signal and reference arms. The noise level goes up as the angular bandwidth increases for the majority of noise sources, while the signal level does not. Consequently, the SNR starts to decrease as we attempt to achieve higher surface density and it becomes necessary to track carefully each of the noise sources and to optimize the system in order to minimize their effect. Through this process we were able to demonstrate a surface density of 10 bits/μm^2 without any errors detected among the bits that were tested.
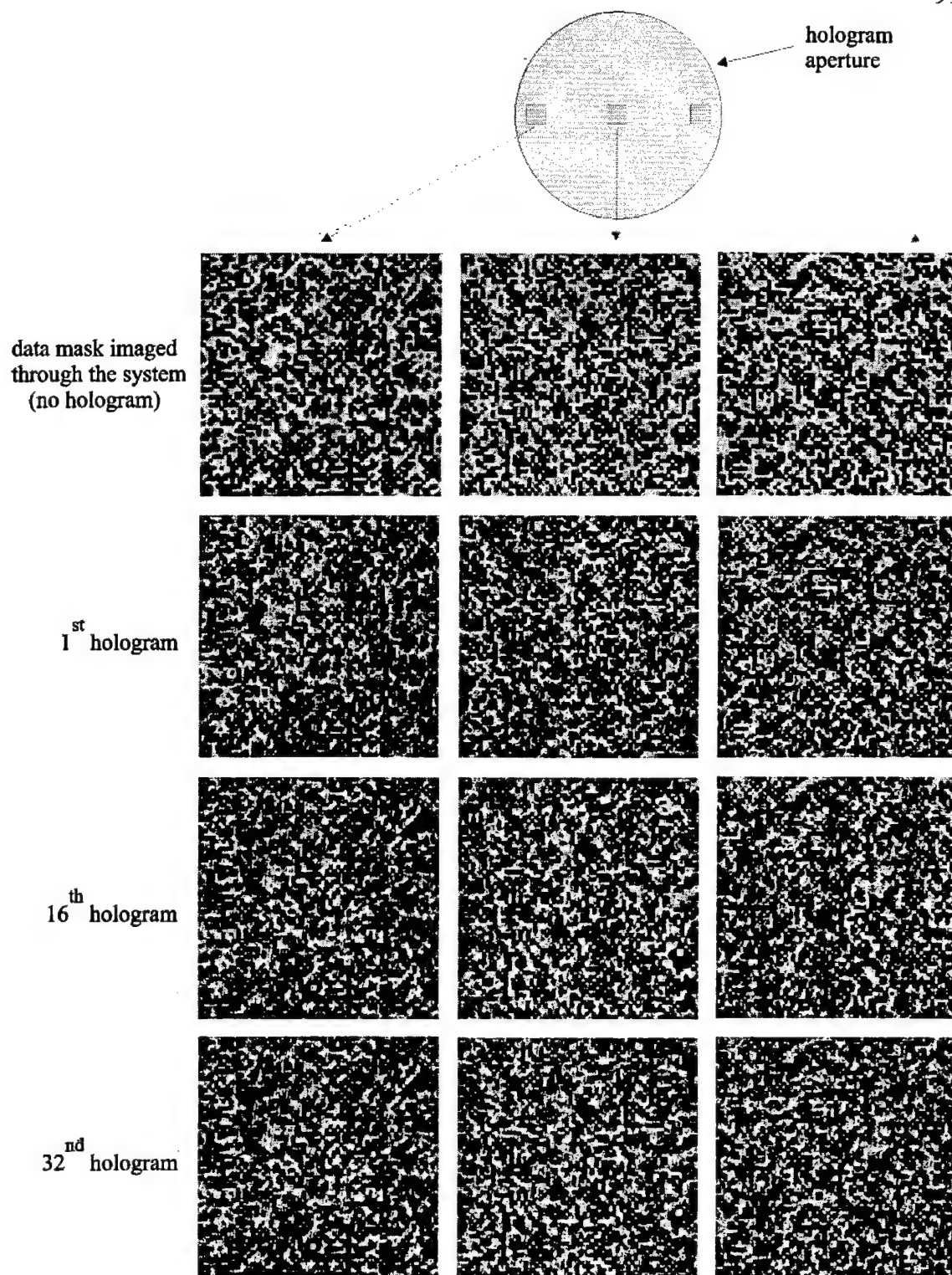
Figure 3.12 : The left, center, and right sides of the data mask and the reconstructions from some of the 32 stored holograms.

We can divide the noise sources into two categories: system noise and hologram noise. Lens aberrations, data mask imperfections, detector noise, scattering, and multiple reflections from lenses and other optical components, laser non-uniformity and fluctuations, and data mask-to-CCD pixel misalignment are examples of system noise. The rest of the noise arises from the hologram itself. Specifically, the hologram can introduce inter-page cross-talk, inter-pixel cross-talk, scattering from the recording material, multiple reflections in the medium, non-uniform diffraction efficiency in the recorded holograms, distortions that are due to surface imperfections, blurring that is due to limited spatial resolution of the material, and material shrinkage. We can separately determine the system noise level by simply imaging a page of data onto the CCD array without any recording material and measure the SNR of the detected image. After minimizing the system noise, we then followed a step-by-step procedure to identify the various sources of hologram noise.

The SNR obtained from various experimental conditions are plotted in Figure 3.13. The first measurement is SNR = 10, which is due to system noise only. We obtained this by transmitting the signal beam through the system without any recording material. The histograms of the 'on' and 'off' pixels were well separated and no errors were observed. Next, a glass substrate on which the photopolymer is usually laminated was placed at the recording plane of the system. The SNR dropped to 9 because of the parallel plate interference effect and the slight aberrations caused by the substrate. When we introduced a piece of UV-cured photopolymer laminated on a glass substrate at the place where the holograms are normally located, the SNR dropped from the upper limit of 9 to approximately 7. The major cause of this drop in SNR is most likely due to internal

reflections within the film and the glass substrate (parallel plate interference effect) and possibly scattering. The interference fringes could perhaps be eliminated with antireflection coatings applied to both the glass and the photopolymer.
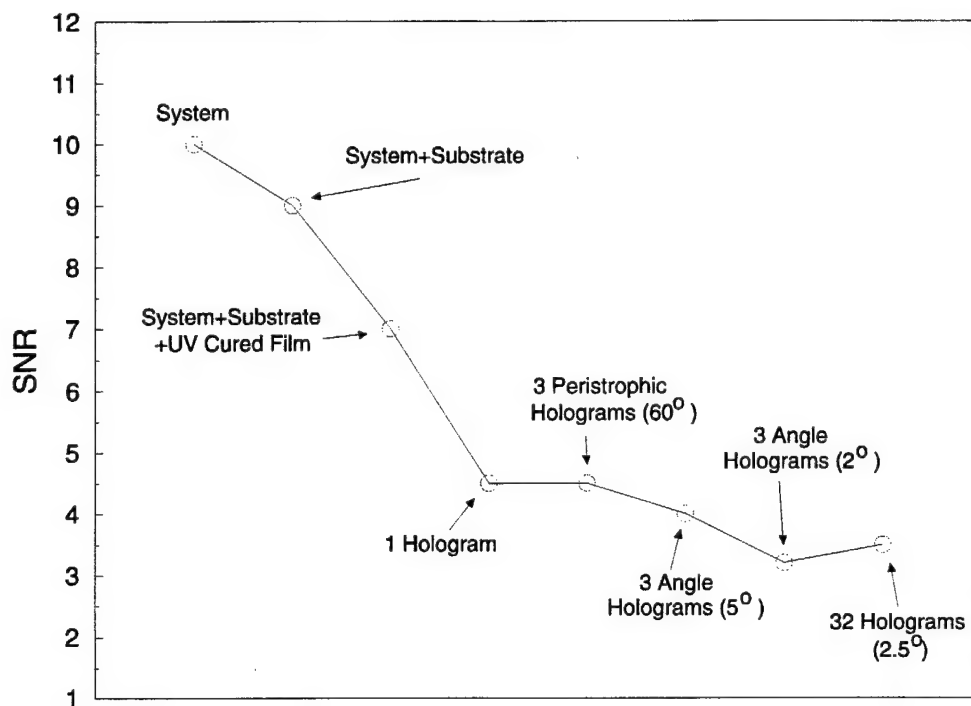

Figure 3.13 : SNR characterization for the high density setup.

A SNR of approximately 4.5 was obtained when a single hologram was recorded and its reconstruction evaluated. The histogram of the 'on' and 'off' pixels were still clearly separated and no errors were observed. This shows that the resolution of the photopolymer is sufficient to record high density holograms with clarity. The reason the SNR dropped from 7 to 4.5 for one hologram is strongly connected to the internal reflections from the boundaries. As stated above, the signal beam creates an interference pattern (fringes) when it passes through the film and the glass substrate. This is also true for the reference beam when it is brought in to record a hologram. Furthermore, when the reference beam is used to reconstruct the stored holograms, it also creates an interference

pattern that can be seen with the naked eye on the film. Therefore, the reconstruction suffers from the fringes that are in the reference beam twice. Other effects such as the fidelity of the recording material also plays a role in degrading the SNR.

When compared to the SNR of a single hologram, the average SNR of three peristrophically multiplexed holograms (60° separation) remained at ~4.5. This indicates two things: peristrophic multiplexing does not introduce cross talk (as expected), and the superposition of three holograms does not significantly deteriorate the quality of each hologram. The average SNR of three angle-multiplexed holograms (5° separation) dropped to ~4. This is a clear indication of cross-talk noise in the angle holograms. The SNR dropped further to ~3.2 for angle holograms separated by 2°. With a SNR of only 3.2, the histograms of the 'on' and 'off' pixels almost overlapped, but no errors were observed. The final SNR for the 10 bits/μm^2 experiment is 3.5 with an angular separation of 2.5° between angle holograms.

In the course of the experiment, we discovered that the strength of the holograms is strongly related to the SNR. If the initial holograms are very strong, then the later holograms will suffer from the index modulation of the previously stored holograms. This effect can be observed by imaging the data mask though the spot where strong holograms have been stored. The image of the data mask gets blurred and the 'off' pixels get filled with some of the 'on' pixels' energy. This means that any subsequent holograms recorded at that spot would have poor SNR. On the other hand, the holograms cannot be recorded too weakly neither. For the high density setup, the hologram area is about 1.77 mm^2 and the final reconstruction at the detector plane has an area of about 12 cm^2. There is nearly a 700X increase in area from the hologram to the detector plane. This means the

intensity of the reconstructed hologram would decrease by the same amount at the detector plane. In order to have good SNR, a sufficient number of photons must strike the detector array. For the 10 bits/μm^2 experiment where 32 holograms were stored, the diffraction efficiency of each hologram was approximately $\eta = 3.5 \times 10^{-3}$. This hologram strength was chosen through trail-and-error to avoid excessive index modulation of later holograms and to have sufficient diffracted power to achieve a good SNR. With the experimentally measure $M/\#$ of 6.5 (Chapter 2.4.4) for the HRF-150-100 photopolymer, the diffraction efficiency per hologram should have been around 4×10^{-2} for each of the 32 holograms. However, we did not use the entire dynamic range of the photopolymer for this experiment since that would have caused index modulation problems. Instead, we used only approximately 30% of the dynamic range of the material, which effectively reduced the $M/\#$ to 1/3rd of its original value.

For the high density experiment, we also determined that with a proper setup, photopolymer scattering is not a big factor in hologram SNR. A proper setup means: (1) the aperture of the reference beam is closed down to the minimum size required so the energy from material scattering does not over-power the reconstruction. (2) The reference beam angle is large enough so the scattering cone have very little energy in the signal beam direction. In the high density setup, we used a 4-f system plus an iris in the reference arm to control the aperture of the reference beam. This allowed us to illuminate only the area of the photopolymer that had holograms stored. To reduce the amount of scattered energy that gets collected by the signal beam lenses, the reference beam angle was made bigger than the signal beam angle by 10° in our setup. The diffraction efficiency of each hologram was approximately 3.5×10^{-3} whereas the equivalent

diffraction efficiency of the light scattered from the photopolymer when it is illuminated by the reference beam is only 3×10^{-5}. Therefore, scattering is a negligible contributor to the overall noise level for the 10 bits/μm^2 experiment.

Even with all the noise sources discussed above, the nine sampled reconstructions from the 32 holograms stored (Figure 3.12) had no detectable errors. The combined histogram from the nine different sampled windows is shown in Figure 3.14. We obtained an estimate for the probability of error by fitting a first-order χ^2 distribution to the histogram of Figure 3.14. The estimated probability of error from this model is approximately 10^{-4}, in the same range as compact discs (10^{-4} to 10^{-6}).
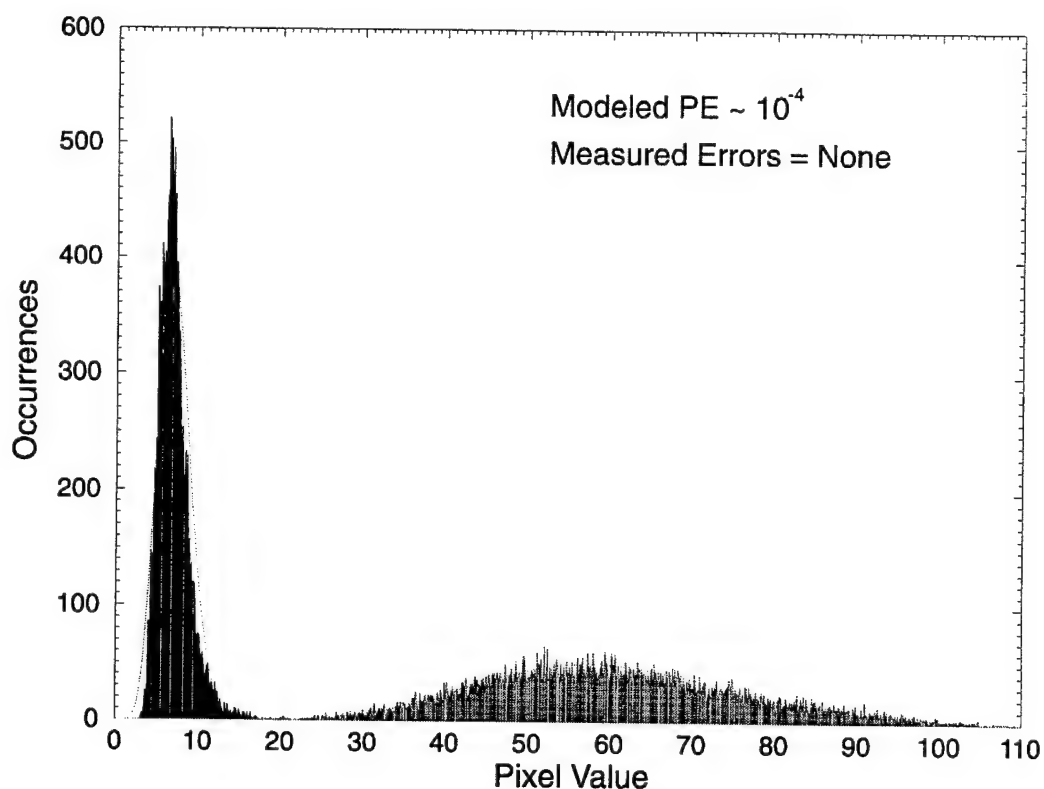

Figure 3.14 : Combined histogram obtained from nine different sampled windows for the 10 bits/μm^2 experiment using angle and peristrophic multiplexing.

3.3 Holographic 3-D Disks Using Shift Multiplexing

3.3.1 Introduction

In the previous experiment (Chapter 3.2), we demonstrated a storage density of 10 bits/μm^2 using the 100 micron thick DuPont photopolymer. This was achieved by utilizing a combination of peristrophic and angle multiplexing to store 32 overlapping holograms in one area. Each hologram contained approximately 590,000 pixels (bits) for a total of 2.25 MBytes stored in an area of 1.77 mm^2. The raw estimated bit-error-rate (BER) of the system, determined from reconstructed holograms, was around 10^{-4}.

The above experiment is important for several reasons. First, it shows that a storage density of approximately ten times higher than the compact disc can be achieved holographically, using a recording material that is only 100 micron thick. The projected storage density of that system, using a 1 mm thick recording material, is around 100 bits/μm^2. Second, it shows that high resolution holograms can be recorded in the DuPont photopolymer with sufficient signal-to-noise ratio for commercial applications. With two-dimensional error correction codes, the BER can be reduced to the industry standard, with overheads similar to CDs.

In this section, we discuss a new 10 bits/μm^2 experiment that builds upon the previous experiment. A novel multiplexing method called shift multiplexing, is used to record holograms in a holographic 3-D disk. The disk architecture is interesting to us because most of today's randomly accessed mass storage devices are based on the disk architecture. The compact disc alone has revolutionized the way information is stored and distributed. Its low cost per megabyte, high storage capacity, and removability makes it ideal for mass distribution. In this section, we will show a holographic 3-D disk system that is very similar to the compact disc in most respects. However, our holographic

system offers higher storage density and potentially greater data transfer rate than conventional disk based storage devices.

3.3.2 Holographic 3-D disks

There are several formats in which holograms can be stored on a disk. Stacks of holograms can be stored in non-overlapping areas on the disk as shown in Figure 3.15. In each stack, a number of holograms are multiplexed by using angle, peristrophic, wavelength, and other multiplexing methods to increase the surface density. To increase the storage capacity of the disk, stacks of holograms are then spatially multiplexed on non-overlapping areas on the disk. This is the most natural method of storing holograms on a disk using conventional multiplexing techniques. During readout, the disk is first rotated to the correct stack and then a reference beam of the proper angle or wavelength is introduced to reconstructed the desired hologram. To accomplish this, a complicated reference beam delivery system or a wavelength tunable laser is required.
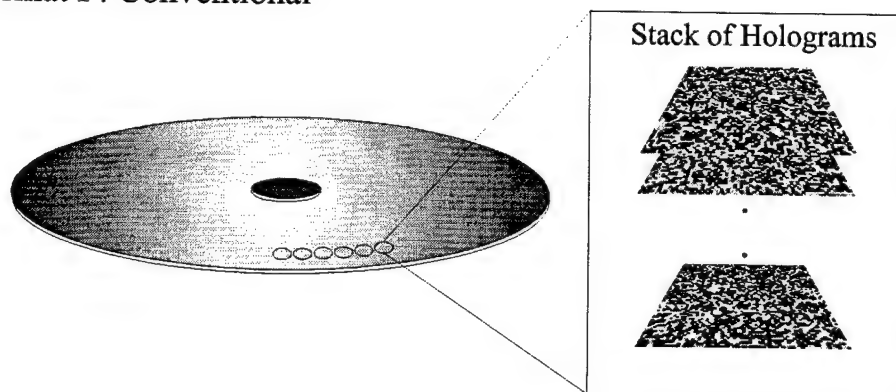
Format I : Conventional

Figure 3.15 : Format in which holograms can be stored in a holographic 3-D disk using conventional multiplexing methods.

For example, Figure 3.16 shows the schematic diagram of a disk system using angle and peristrophic multiplexing. The signal beam is on axis and the reference beam is off axis (with respect to the disk's surface normal). The angle of the reference beam is changed to achieve angle multiplexing while the rotation of the reference beam around the signal beam gives peristrophic multiplexing. The disk motion provides access to different spatial locations for storage and the entire signal and reference beam assembly can be moved radially to access more locations. Using this setup, we can pretty much repeat the same high density experiment of Chapter 3.2 at many locations on the disk. The only problem is how to deliver the reference beam in the angles shown in Figure 3.16.
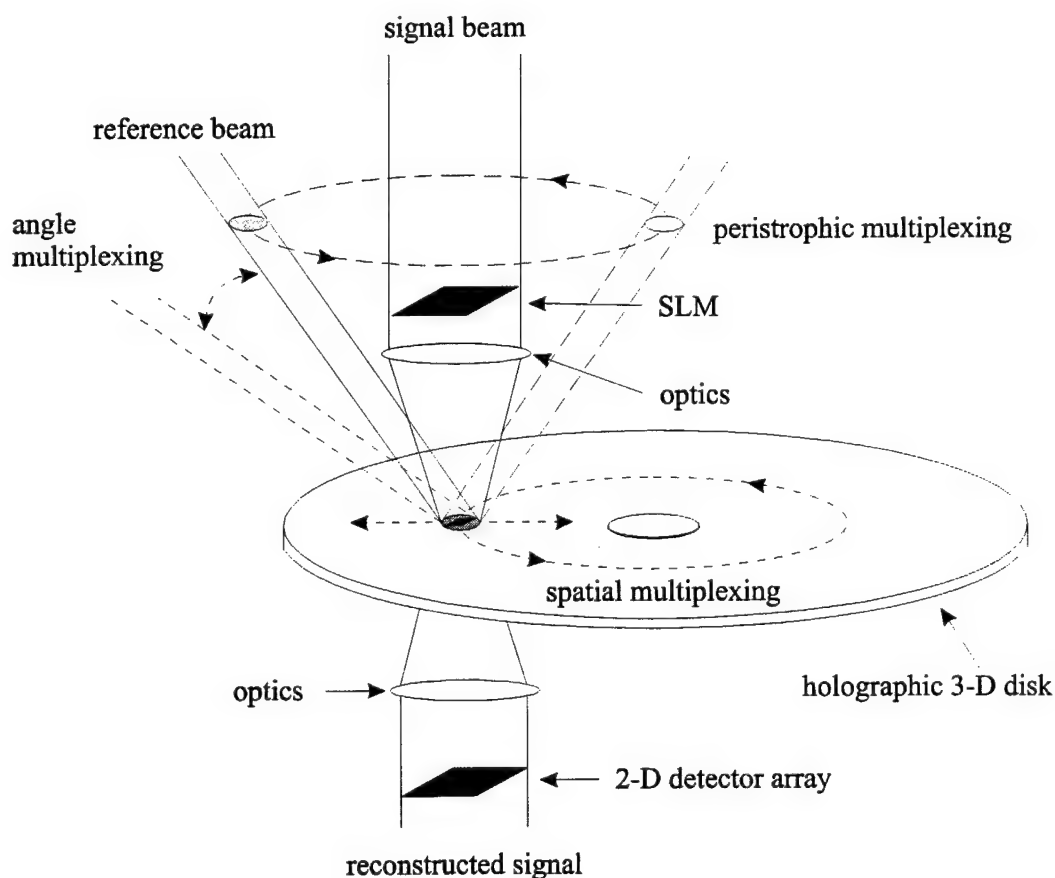
Figure 3.16 : Schematic diagram of a holographic 3-D disk system using peristrophic and angle multiplexing.

Furthermore, holograms stored in this stack format is unnatural when compared to the compact disc. It is highly desirable in the disk architecture, to store holograms sequentially in a spiral fashion, from the inner radius to the outer. Likewise, the reference beam system should be as simple and practical as possible. To accomplish these two goals, we use a new multiplexing method called shift multiplexing in our holographic 3-D disk system. Shift multiplexing allows holograms to be stored sequentially in partially overlapping areas as shown in Figure 3.17. The spiral tracks can also be overlapped to further increase the surface density. With shift multiplexing, holograms are recorded with a multiple planewaves reference beam or a spherical reference beam. When a hologram recorded with a multiple planewaves reference beam is shifted, reconstructions from the multiple planewaves will add destructively and constructively in a periodic cycle. Overlapping holograms can be stored at the nulls of the periodic function. When a hologram recorded with a spherical reference beam is shifted slightly, the wavefront the hologram experiences becomes different from the wavefront used to record it. Therefore, the hologram will simply Bragg mis-match (if shifted parallel to the plane of interaction) or the reconstruction will deflect at a different angle and miss the detector (if shifted perpendicular to the plane of interaction). Hence, instead of using several different reference beam angles or wavelengths to multiplex holograms in a stack, a single fixed reference beam can be used to record partially overlapping holograms in a spiral track. During readout, the disk is spun and sequential holograms will become Bragg matched and be readout by the fixed reference beam, much like the pits on the compact disc.

However, with our holographic system, instead of reading out one bit at a time as the disk rotates, an entire page of a hologram can be read out at a much higher data transfer rate. Furthermore, since holograms overlap in area, the surface density is increased. The theory behind shift multiplexing is discussed in the next section.

Format II : Shift
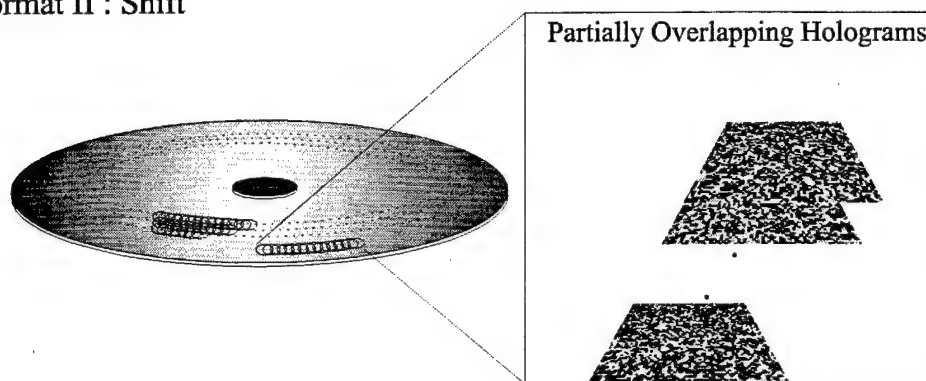
(b)

Figure 3.17 : Format in which holograms can be stored in a holographic 3-D disk using shift multiplexing.

3.3.3 Shift Multiplexing

3.3.3.1 Multiple Planewaves Reference Beam

In all the experiments shown previously in this thesis, the holograms were recorded by interfering the signal with a planewave reference beam. By changing the angle of the reference beam (either angularly or peristrophically), more holograms can be multiplexed at the same location. For the holographic 3-D disk, we would like to store holograms sequentially by slightly shifting the recording material between holograms. This slight displacement must somehow get rid of the reconstructions from the previously stored holograms, so that another hologram could be stored without too much cross-talk. For a hologram recorded with a single planewave reference beam, shifting the recording

material during reconstruction slowly decreases the hologram's diffracted power. This is because part of the hologram's area moves out of the reference beam's illumination (Figure 3.18). The amount of shift required for the hologram's diffracted power to reach zero is the width of the hologram, x. This means that in order to store another hologram, we would have to shift to a new location on the recording material. Therefore, when using a single planewave reference beam, no more than one hologram can be stored at a location by shifting the recording material.
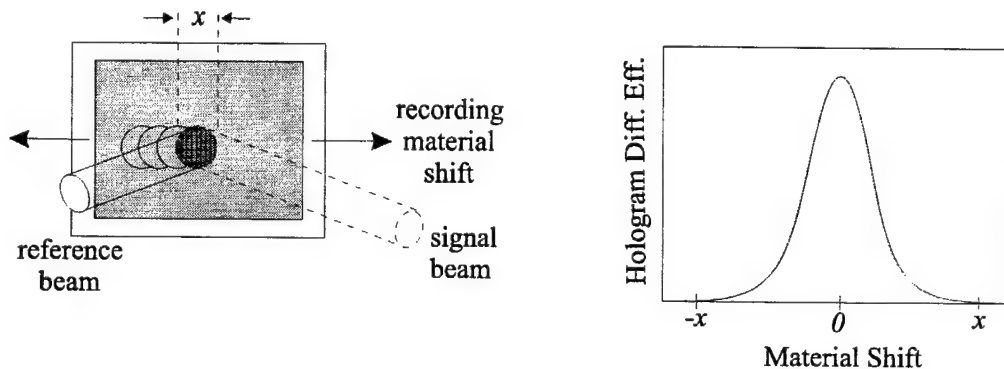

Figure 3.18 : Shift multiplexing using a single planewave reference beam.

What if instead of recording with only one planewave reference beam, a hologram was recorded by interfering the signal with multiple planewaves simultaneously? In that case, each planewave will form a separate hologram with the signal beam. Initially, all the reconstructions from the separate holograms will be in phase and add constructively. However, if the recording material is shifted during reconstruction, the phases of the holograms will start to differ and with sufficient shift, the holograms will completely destructively interfere and a null in diffracted power appears (Figure 3.19). The reconstruction's diffracted power is periodic because at some point in shift, the holograms will again add constructively. However, the diffracted power will not be as strong as at

the zero shift position because some of the hologram's area is outside the reference beam's illumination.
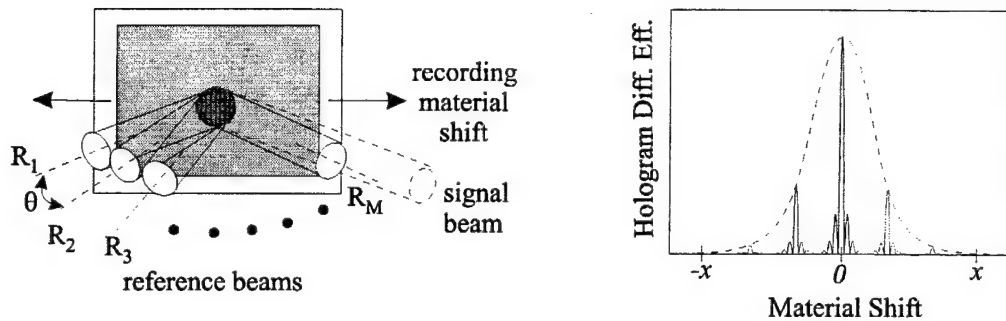

Figure 3.19 : Shift multiplexing using multiple planewaves.

The theory behind shift multiplexing using multiple planewaves is very simple. Figure 3.20 (a) shows a diagram of multiple planewaves recording a hologram with the signal beam. The angular separation between the reference planewaves is θ and they all have the same intensity. Each planewave records a separate hologram with the signal beam and all the holograms are in-phase when the recording material is not shifted. For the reference beam that makes an angle of $\theta/2$ with respect to the surface normal of the recording material (Figure 3.20 (b)), shifting the recording material during reconstruction introduces a phase difference of :

$$\frac{2\pi}{\lambda} \delta \sin(\frac{\theta}{2}) = \frac{2\pi}{\lambda} \zeta \qquad (3.9)$$

where δ is the shift distance, ζ is the phase introduced, and λ is the wavelength of the laser. Since the reference planewaves are symmetric with respect to the surface normal of the recording material, the sum of the phases for all the shifted holograms can be written as :

$$\text{sum of the phases} = \sum_{\theta_m} e^{i\frac{2\pi}{\lambda}\delta \sin(\theta_m)} \qquad (3.10)$$

where $\theta_m = \left(m - \dfrac{M-1}{2} \right)\theta$, for $m = 0,..., M - 1$, and M is the total number of planewaves

in the reference arm. Eq. 3.10 can be simplified further to show the intensity of the

reconstructed signal as a function of recording material shift δ:

$$I(\delta) = \left(\sum_{\theta_m} 2\cos(\frac{2\pi}{\lambda} \delta \sin(\theta_m)) \right)^2 \tag{3.11}$$

Assuming the angle between the reference planewaves is small ($\theta \simeq \sin\theta$), the period of

Eq. 3.11 is :

$$\delta_P = \frac{\lambda}{\theta} \tag{3.12}$$

and the zeros of Eq. 3.11 within a single period are at :

$$\delta_z = \frac{\lambda}{M\theta} l, \qquad l = 1, \ldots, M\text{-}1 \tag{3.13}$$
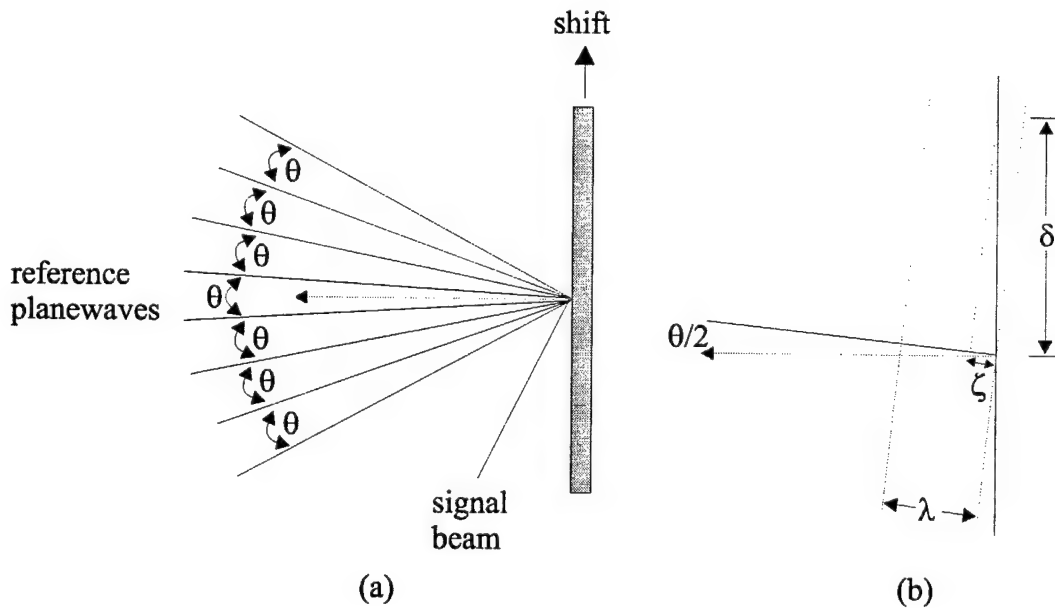

Figure 3.20 : Multiple planewaves shift multiplexing setup. (a) The delivery of multiple planewaves. (b) Phase introduced when shifted, for one planewave.

According to Eq. 3.13, it is theoretically possible to record M partially overlapping holograms within the period defined by Eq. 3.12, each separated by $\delta = \lambda/M\theta$. After recording the first hologram, there are M-1 nulls before the first hologram becomes bright again. At each of the nulls, one additional hologram can be written, bring the total number of holograms stored to M. To store more holograms, the material must be shifted until the first hologram is completely out of the illumination of the reference beam, eliminating the periodicity of the first hologram.

M, the number of planewaves in the reference arm determines how many overlapping holograms could be stored in one period and θ, the angle between the reference planewaves determines the period of the stored hologram. A large M allows many holograms to be stored in a given area and a small θ increases the period width. So far, the thickness of the recording material, the signal beam angle with respect to the recording material's surface normal, and the signal bandwidth have nothing to do with the number of holograms that can be stored in a given volume. However, they must be accounted for somehow or else we could break the V/λ^3 density maximum by using a large number of planewaves in the reference arm to store many high bandwidth holograms in a very thin recording material. Figure 3.21 shows the κ-sphere representation of a two planewaves reference beam recording a hologram with a planewave signal beam. Each planewave reference records a separate grating with the signal beam as shown in Figure 3.21 (a). Since the grating vectors can be read by either reference beams, the κ-sphere diagram during reconstruction looks like Figure 3.21 (b). Figure 3.21 (b) shows that two 'ghost' signals get reconstructed along with the desired signal beam. If the angle between the two planewaves reference is θ, then the angles the

'ghost' signals make with the desired signal is $\pm\theta$. The strength of the 'ghost' signals depends on the Bragg selectivity since the grating vectors for them are off the κ-sphere. There are two ways to get rid of these 'ghost' images. One is to select the correct angle between the reference planewaves so the 'ghost' reconstructions are at the nulls of the Bragg selectivity. This can be achieved by picking a minimum $\theta \simeq \lambda/L\tan\theta_s$, where λ is the wavelength, L is the thickness of the recording material, and θ_s is the angle of the signal beam with respect to the surface normal of the recording material. The second method is to filter the 'ghost' reconstructions with a spatial filter placed somewhere before the detector. In this case, the minimum angular separation between the reference planewaves for thin recording materials is the bandwidth of the signal beam. Both methods restrict the minimum angular separation between the reference planewaves. This limits the number of reference planewaves that can be in the reference arm and therefore the number of shift holograms that could be stored in a setup.
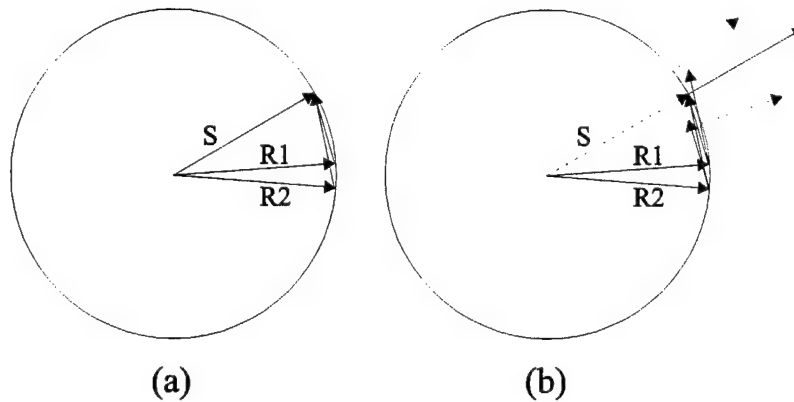

(a)   (b)

Figure 3.21 : κ-sphere representation of a hologram recorded with a two planewaves reference beam. (a) Double grating formation during recording. (b) Reconstruction showing 'ghost' images.

To experimentally demonstrate shifting multiplexing with a multiple planewaves reference, we constructed the setup shown in Figure 3.22. A multiple planewaves

reference was generated by recording 20 regularly spaced planewave holograms in a piece of HRF-150-38 photopolymer. The reference planewave* was fixed while the angle of the signal planewave* was changed by 0.25° between holograms with a rotational stage and a 4-f lens system. The diffraction efficiency of the 20 planewave holograms were equalized by using a recording schedule generated with the iterative method shown in Chapter 2.4.3. Once the holograms have been stored and the photopolymer properly cured, illumination with the fixed reference planewave* reconstructs all 20 planewave holograms simultaneously. The second 4-f system re-images the fan of 20 planewaves onto another piece of HRF-150-38 photopolymer as the multiple planewaves reference beam. The second photopolymer was mounted on a linear translation stage so the diffracted power from the stored holograms could be measured as a function of material displacement. Another signal planewave was bought in to record a hologram in the photopolymer with the multiple planewaves reference. For this experiment, the angular separation of 0.25° was much smaller than the Bragg selectivity of the 38 micron thick photopolymer. Therefore, a spatial filter was placed at the Fourier plane of the third 4-f system to block the 'ghost' images. A power detector was placed at the image plane of the third 4-f system to measure the diffracted power of the reconstructed hologram.

Figure 3.23 shows the diffracted power of a planewave hologram recorded using the setup shown in Figure 3.22 as the photopolymer was shifted. The diffracted power as a function of material shift is a periodic function modulated by the width of the hologram. Shifting the hologram moves some of the hologram's area out of illumination and therefore, the intensity of the periodic function falls off as the material is shifted further away from the center. The blow-up of the center portion of Figure 3.23 is shown in

Figure 3.24, along with the theoretical prediction according to Eq. 3.11. The experimentally obtained period of 115 μm and a null spacing of 5.75 μm match closely with the theoretical prediction of $\delta_P = 112$ μm and $\delta_Z = 5.6$ μm. The slight discrepancy is caused by the $\theta \simeq \sin\theta$ approximation used in the theoretical prediction.
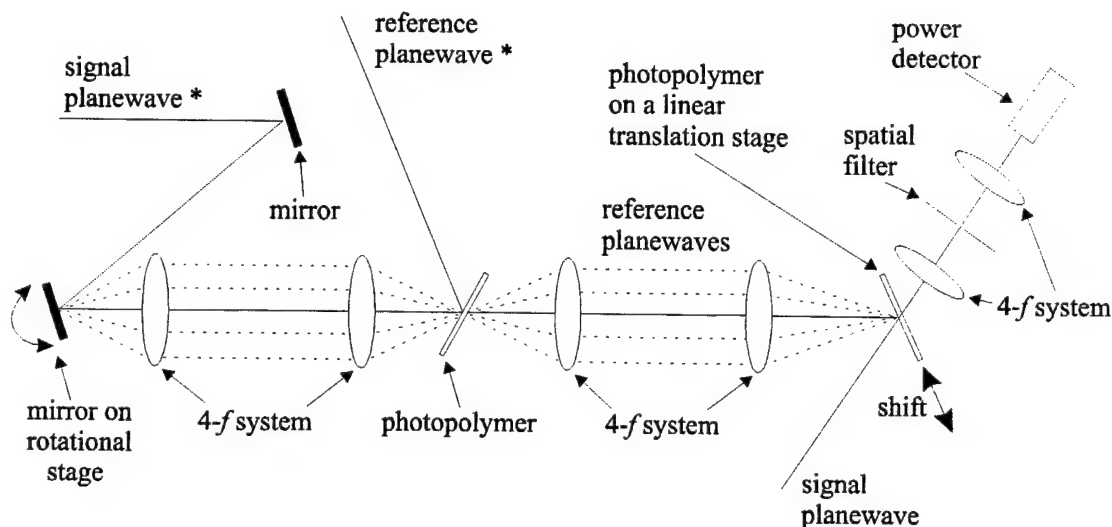

Figure 3.22 : The setup used to experimentally demonstrate shift multiplexing using a multiple planewaves reference beam.
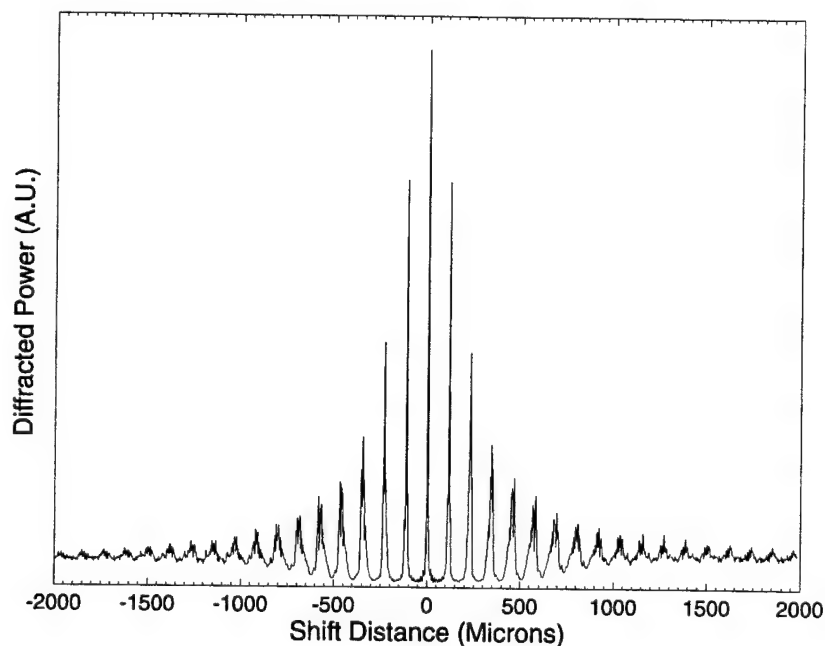

Figure 3.23 : The diffracted power as a function of material shift for a planewave hologram recorded with a 20 planewaves reference arm.
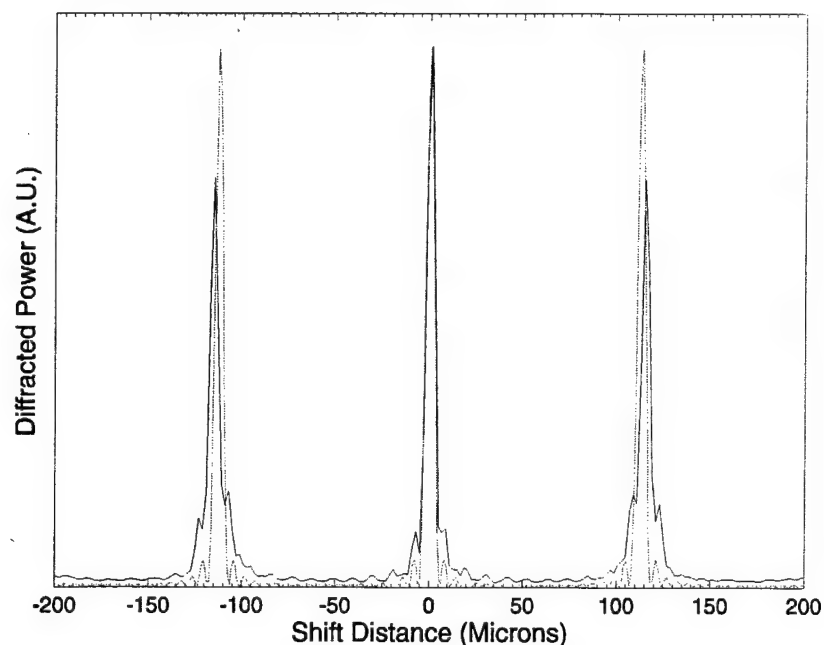
Figure 3.24 : Blow-up of the center portion of Figure 3.23 and the theoretical intensity curve according to Eq. 3.11.

Figure 3.25 shows the diffracted power a 100×100 random bit pattern image plane hologram recorded and measured with the same setup. The period and the null spacing remained the same but the nulls are less distinguishable due to the increased signal beam bandwidth. As with angle multiplexing, the nulls rarely have zero diffracted power in an actual experiment. In order to reduce cross-talk noise, angle multiplexed holograms are usually stored several nulls away from one another. Figure 3.26 shows three holograms shift multiplexed within one period. Holograms #2 and #3 were multiplexed within a single period of Hologram #1, with comfortable separations. Each hologram have its own array function which decays with material displacement. The quality of the reconstructed holograms were very good with no noticeable cross-talk noise. However from Figure 3.26 it is clear that even though we used 20 planewaves in

the reference arm, fewer than 20 holograms could be stored within the same period with
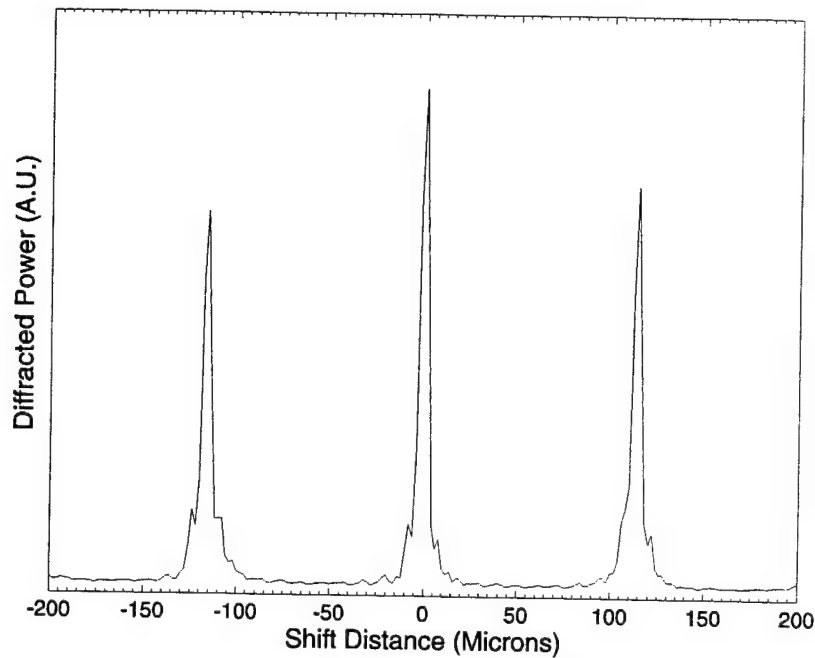
acceptable SNR.

Figure 3.25 : The diffracted power as a function of material shift for a 100×100 random bit pattern image plane hologram recorded with a 20 planewaves reference arm.
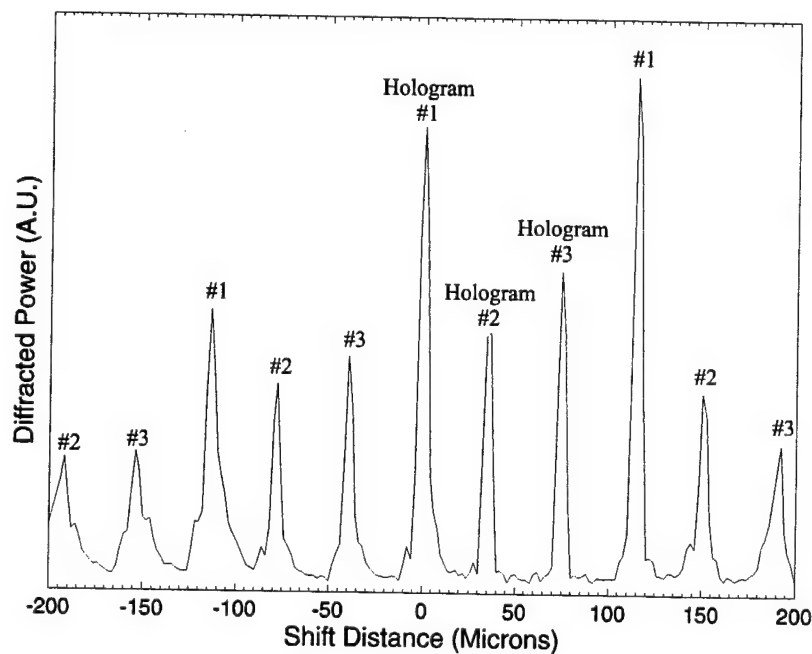

Figure 3.26 : The diffracted power of three random bit pattern image plane holograms recorded with a 20 planewaves reference beam as a function of material shift.

According to Eq. 3.12 and 3.13, by decreasing θ and increasing M, it is theoretically possible to store holograms that are very sharp (null spacing is very small) and have no periodicity (the period is larger than the width of the hologram). To demonstrate this effect, we replaced the photopolymer that generated the multiple planewaves with an etched glass plate. The glass plate was patterned with random on and off lines with a pitch of 5 μm and then etched to roughly $\lambda/2$ in depth. When illuminated on axis, the etched glass plate produced a continuous fan of planewaves in the direction that is perpendicular to the etched lines. The second 4-f system in the setup shown in Figure 3.22 re-images the continuous fan of planewaves onto the recording material as the multiple planewaves reference beam. Figure 3.27 shows the diffracted power as a function of material shift for a planewave hologram recorded with a continuous fan of planewaves. The null spacing became much closer together and the periodicity of the hologram was destroyed. However, since the angular separation between the reference planewaves was very small, there was no way to filter all the 'ghost' reconstructions with the spatial filter at the Fourier plane of the third 4-f system. Therefore the noise floor is much higher. This means that even though the holograms can be multiplexed closer together using the continuous fan of planewaves, the increase in noise will limit the maximum number of holograms that could be stored in one area.

So far we have only talked about generating a fan of planewaves in one-dimension to store shift multiplexed hologram. The direction of shift is the same as the spread of the planewaves. We could also generate a 2-D fan of planewaves to store shift multiplexed holograms in two orthogonal directions. This further increases the number of holograms

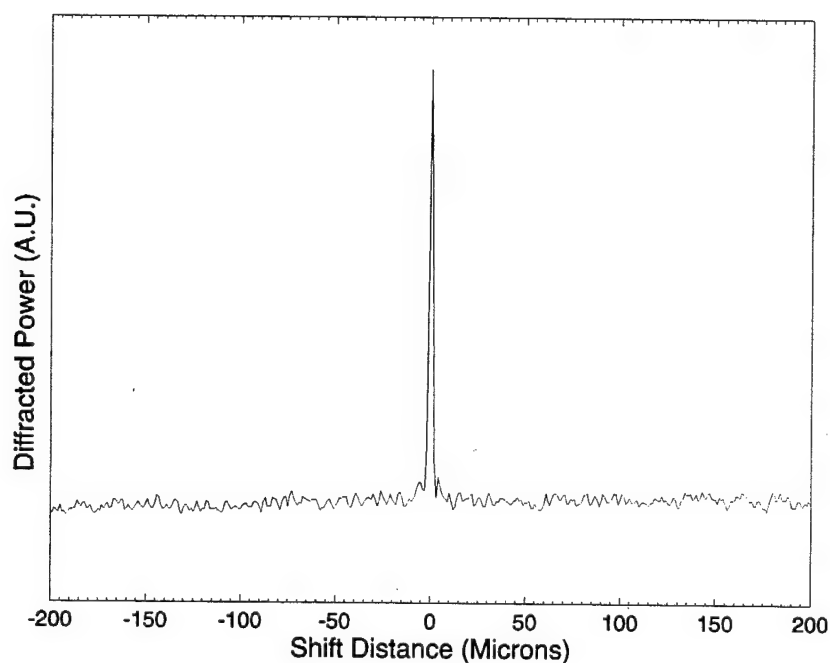that could be stored in one area, which translates into an increase in surface density and storage capacity.

Figure 3.27 : The diffracted power as a function of material shift for a hologram recorded with a continuous fan of planewaves.

3.3.3.2 Spherical Reference Beam

In the previous section, when the angular separation between the planewaves was decreased and the number of planewaves was increased (in the reference arm), the one-dimensional multiple planewaves reference beam looked a lot like a cylindrical beam. Perhaps we could use a cylindrical lens to generate a continuous fan of planewaves for shift multiplexing. However, a cylindrical lens when used as shown in Figure 3.28 does not actually produce a fan of planewaves, but a fan of rays. These rays are spatially separated and they converge to a line at the focal plane and then diverge again. When a hologram is recorded with the setup shown in Figure 3.28, the diffracted power still

decreases rapidly as the recording material is shifted parallel to the fan of rays, but the mechanism is different when compared to holograms recorded with a fan of planewaves.
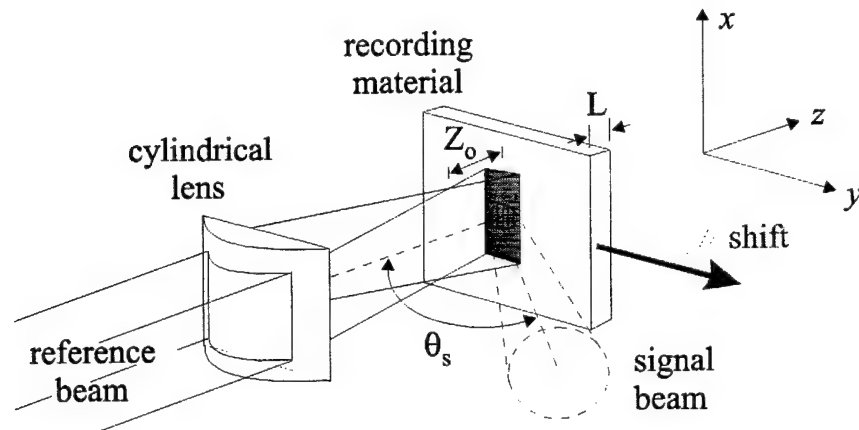

Figure 3.28 : Using a cylindrical lens to produce a fan of rays for shift multiplexing.

Figure 3.29 (a) shows a cylindrical reference beam recording a hologram with a planewave signal beam in the y-z plane of Figure 3.28. During grating formation, the hologram is formed with the rays in the reference beam as shown in Figure 3.29 (a). If the recording material is shifted during reconstruction (Figure 3.29 (b)), the angle of the rays the hologram now experiences is different from the rays used to record it. Therefore, the hologram will Bragg-mismatch very similar to the angle multiplexing effect.
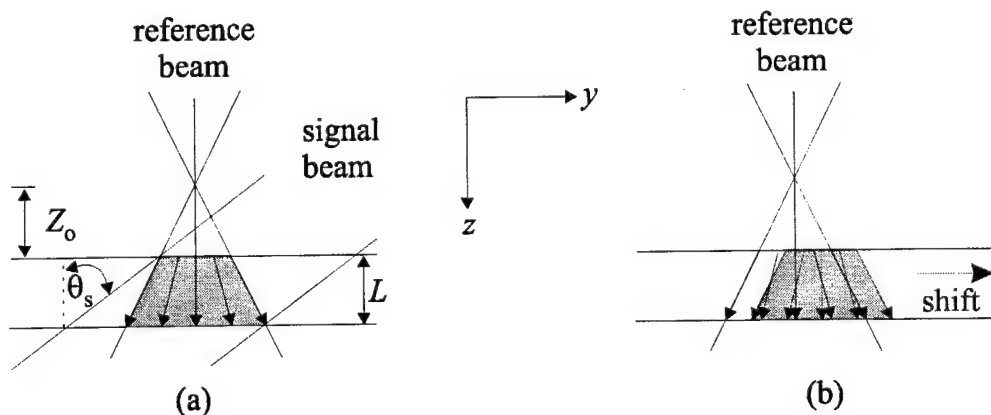

Figure 3.29 : Cylindrical reference beam shift multiplexing mechanism. (a) A cylindrical reference beam recording a hologram. (b) Shifting the recording material during reconstruction.

For thin recording materials such as DuPont's photopolymers, the amount of shift required to reach a null in the Bragg condition can be computed easily. For the setup shown in Figure 3.29 (a), the center ray of the cylindrical reference beam is on-axis and the signal beam is off-axis by θ_s, both measured with respect to the surface normal of the recording material. The focal plane of the cylindrical beam is Z_o away from the recording material and the recording material has a thickness of L. For the hologram formed between the on-axis reference ray and the off-axis signal beam, the amount of angular change introduced by shifting the recording material a distance δ is $\Delta\theta \simeq Z_o\delta$. According to Bragg theory, the amount of angular change required to reach the first null is given by Eq. 3.7. For the case where $\theta_r = 0°$, Eq. 3.7 can be simplified to :

$$\Delta\theta = \frac{\lambda}{L \tan \theta_s} \tag{3.14}$$

By equating the angular change to the angular selectivity, a shift distance of $\delta_{\text{Bragg}} = \lambda Z_o/L\tan\theta_s$ is required to reach the first null. As a fudge factor to prevent δ from going to zero when the cylindrical reference beam is focused on the recording material ($Z_o = 0$), we need to model the finite numerical aperture (NA) of the cylindrical reference lens. The finite numerical aperture broadens the selectivity curve by an additional $\delta_{\text{NA}} \simeq \lambda/2(\text{NA})$. For most setups, δ_{NA} is extremely small when compared to δ_{Bragg}. The final equation for the amount of shift required to reach the first null of the Bragg condition for a hologram recorded between the on-axis reference ray and a planewave signal with an angle of θ_s is :

$$\delta = \delta_{\text{Bragg}} + \delta_{\text{NA}} \cong \frac{\lambda Z_o}{L \tan \theta_s} + \frac{\lambda}{2\,\text{NA}} \tag{3.15}$$

Eq. 3.15 only predicts the null spacing of the grating recorded between the on-axis reference ray and a planewave signal with an angle of θ_s. In order to get the selectivity function for the entire hologram, it is necessary to compute the diffraction pattern for all the gratings produced by the different rays in the reference beam and then sum them together. This is demonstrated later for a special recording geometry (Chapter 3.3.3.3). Nevertheless, Eq. 3.15 tells us how the shift selectivity function δ, depends on λ, Z_o, θ, and L. As with angle multiplexing, shorter wavelengths, thicker recording materials, and larger angles between the signal and reference beams improve the selectivity function. The extra parameter, Z_o, determines the rate of angular change as a function of material shift. A small Z_o causes the ray angle to change rapidly and therefore, improves the selectivity. The minimum Z_o that can be used in a setup depends on the spot size of the signal beam on the recording material and the NA of the cylindrical reference beam. The focal plane of the cylindrical reference beam must be placed far enough from the recording material so that the entire signal beam spot is covered by reference illumination during recording. There was no such restriction with the multiple planewaves reference beam setup shown in Chapter 3.3.3.1. The aperture of the multiple planewaves reference beam at the recording material is the same as the aperture of the diffractive element that generated the multiple planewaves. The minimization of Z_o becomes tougher for thicker recording materials because the signal beam defocus inside the material and causes the entrance spot size to be different from the exit spot size. For best quality holograms, Z_o should be chosen to be large enough to cover both the entrance and exit signal spots. So for thick recording materials, there is some cross relationship between Z_o and L. For the thin photopolymers, we can assume Z_o and L to be independent.

There are other differences between using a fan of planewaves and a fan of rays to record shift multiplexed holograms. For holograms recorded with a cylindrical reference beam, the holograms have no periodicity as the recording material is shifted and there are no 'ghost' images. This is because each ray of the cylindrical reference beam illuminates a separate, infinitely small portion of the entire hologram. Each grating vector on the κ-sphere is illuminated by only one reference ray at a time. Therefore, there is no constructive and destructive periodicity and no 'ghost' images.

To further increase the number of holograms stored in a given area, we would also like to shift multiplex in directions other than just parallel to the fan of rays. For example, shift multiplex by moving the recording material in the direction that is orthogonal to the fan of rays. However for a cylindrical reference beam, the holograms experience no change in the ray angles when shifted orthogonally to the fan of rays. Therefore, the holograms would have to be shifted completely out of the reference illumination for the reconstruction to disappear. To introduce a fan of rays in the orthogonal direction, we could use another cylindrical lens to focus the rays in the x-direction but why stop there? Why not introduce a fan of rays in every direction so we could shift in any direction and get some selectivity. A fan of rays in all directions can be generated by using a single spherical lens. Figure 3.30 shows a shift multiplexing setup using a spherical reference lens. Later in Chapters 3.4 and 3.5, we will use a spherical reference beam to record many high density holograms in the same volume.

When the recording material is shifted parallel to the plane of interaction (y-z plane) for a hologram recorded with a spherical reference beam, the response of the stored holograms is exactly the same as the analysis we did for the cylindrical beam

setup. How about when the recording material is shifted perpendicular to the plane of interaction? Figure 3.31 (a) shows a spherical reference beam recording a hologram with a planewave signal beam in the x-z plane of Figure 3.30. The solid lines inside the hologram volume represent the grating and the dotted lines represent the signal planewave. The focal point of the spherical reference beam is Z_o away from the recording material. Figure 3.31 (b) shows the recording material shifted along the x-axis. When shifting orthogonal to the plane of interaction, the holograms are much harder to Bragg-mismatch compared to shifting parallel to the plane of interaction. Similar to peristrophic multiplexing, the initial tendency of the shifted hologram (along x) is to reconstruct at a different angle. Therefore, we call the orthogonal direction the Bragg degenerate direction.
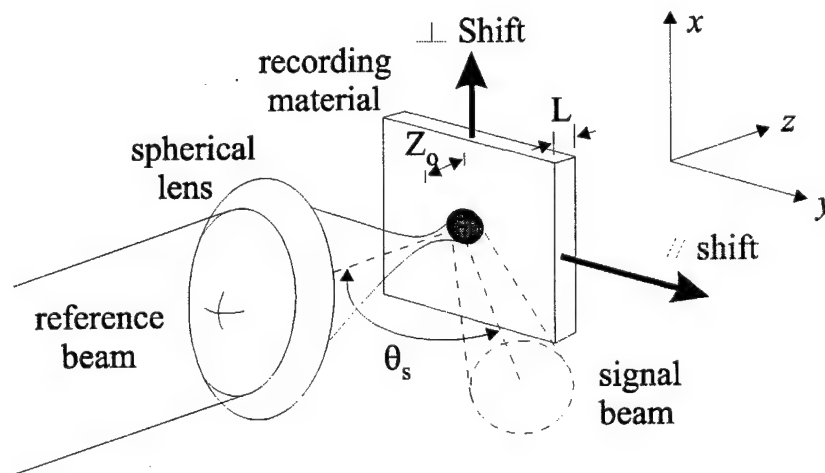

Figure 3.30 : Shift multiplexing setup using a spherical reference beam.
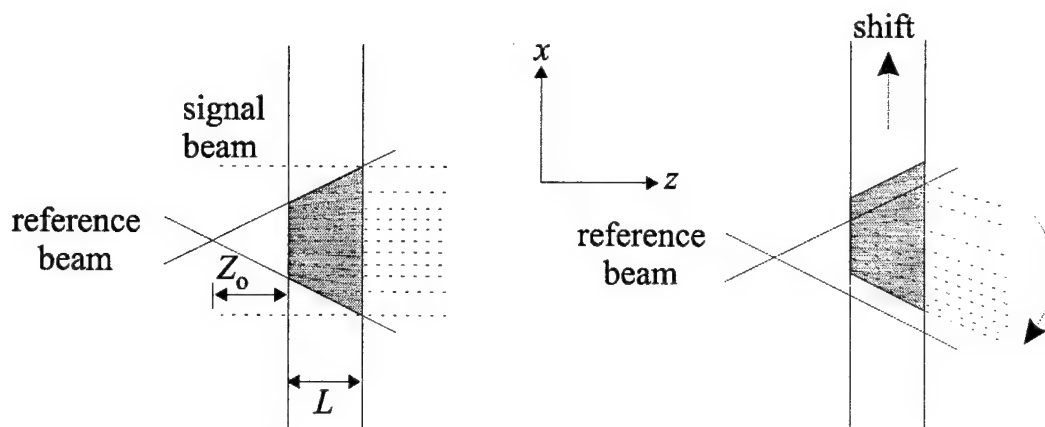
Figure 3.31 : Spherical reference beam shift multiplexing mechanism in the Bragg degenerate direction. (a) A spherical reference beam recording a hologram with a planewave signal beam. (b) Shifting the recording material in the degenerate direction during reconstruction.

Figure 3.32 shows the κ-sphere representation of shift multiplexing using a spherical reference beam. In Figure 3.32 (a), the 2-D fan of rays in the spherical reference beam is represented by a circular patch on the κ-sphere. Each ray on the patch records a separate grating with the planewave signal beam. One such grating is shown in Figure 3.32 for the on-axis reference ray. If the recording material is shifted orthogonal to the plane of interaction, all the recorded gratings move and some get illuminated by rays with different angles and some move out of the reference beam's illumination. However instead of shifting the gratings along the y-axis, which would have immediately caused some Bragg-mismatch (the tip of the grating vectors would either poke through or dive below the κ-sphere), the gratings are moved along the x-axis. For a wide-range of orthogonal shifts, the grating vectors that have no x-components remain relatively on the κ-sphere (for example, the grating vector shown in Figure 3.32). To see this, you have to imagine Figure 3.32 (b) as half of a sphere sticking out of the page and see the movement of the gratings on the sphere as they are shifted along the x-axis. For the grating vectors with some x-components, they reside below the apex of the sphere. Therefore, when

shifted orthogonally, they either initially lift off or dive below the κ-sphere and Bragg-mismatch quickly. In experiments, what is usually seen at the detector plane as the recording material is shifted orthogonally is a strip of Bragg-matched gratings that moves up or down until the reconstruction is out of the field of view of the detector.

In order to store another hologram in the orthogonal direction, the existing hologram must be shifted enough so that either: (1) the grating vectors that have no x-components are shifted completely out of the illumination of the reference beam; (2) the change in the hologram's reconstruction angle is sufficient for filtering, like peristrophic multiplexing; (3) the reconstructions from the grating vectors that have no x-components are shifted off of the detector array. For the first case, the amount of orthogonal shift required is usually half the hologram height, if the reference spot size is the same as the signal spot size on the recording material. For the second case, the amount of orthogonal shift required to move the entire reconstruction out of the detector array's aperture is $\delta \simeq 2Z_o \tan^{-1}(1/2f/\#)$, where Z_o is the distance from the focal point of the spherical reference beam to the recording material and $f/\#$ is the distance between the detector array and the recording material divided by the aperture of the detector array, if free space propagation. If lenses are used in the signal arm, then $f/\#$ is defined as the focal length of the imaging lens divided by the aperture of the detector array. For the last case, the amount of orthogonal shift required to shift just the reconstructions from the grating vectors that have no x-components off of the detector array is $\delta \simeq Z_o \tan^{-1}(1/2f/\#)$ (half the hologram height if the signal and reference beams have the same bandwidth and overlap exactly). The last case is the most common since we usually only need to shift the strip of Bragg match reconstructions off the detector array in order to store another hologram.
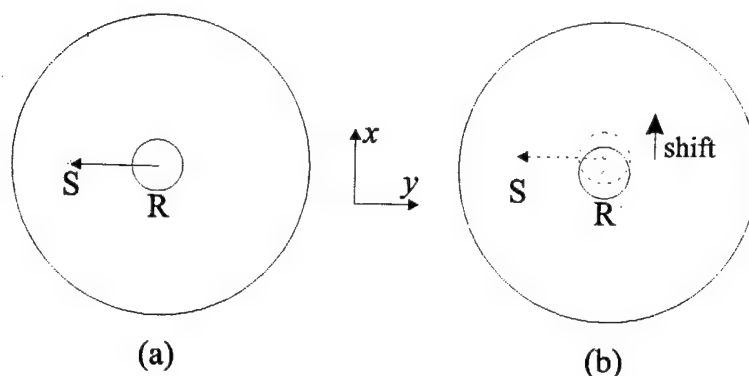
Figure 3.32 : κ-sphere representation of shift multiplexing using a spherical reference beam. (a) Recording a hologram between a planewave signal beam and a spherical reference beam. (b) Shifting in the degenerate direction during reconstruction.

In the high density experiment shown in Chapter 3.2, we were able to read only a portion of the reconstructed hologram at a time due to photopolymer shrinkage. Different portions of the hologram can be obtained by slightly de-tuning the angle of the reference beam. For a hologram recorded with a spherical reference beam, it is not clear how one would re-Bragg match a shrunken hologram. In the next chapter we discuss an off axis recording geometry, using a spherical reference beam, that eliminates the effect of photopolymer shrinkage on hologram reconstruction. We will continue our analysis of spherical reference beam shift multiplexing for this special geometry there.

3.3.3.3 Shrinkage Insensitive Recording Geometry

For the high density holograms recorded with a planewave reference beam shown in Chapter 3.2, portions of the recorded gratings would lift-off or dive below the κ-sphere due to shrinkage. These Bragg-mismatched portions can be brought back on the κ-sphere by changing the reference beam angle slightly during reconstruction. However, the reconstruction angle of the re-Bragg-matched portion is different when compared to the

original signal direction and only a portion of the entire hologram can be Bragg-matched at a time. It might be possible to use a slightly cylindrical reference beam to re-Bragg-match the entire hologram at the same time. However, the reconstructed image would still be a little bit compressed.

For a high density hologram recorded with a spherical reference beam using the setup shown in Figure 3.30, to re-Bragg-match a portion of the shrunken hologram, we need to both shift the recording material and move the focus of the reference beam lens. Shifting the recording material brings some grating vectors back on the k-sphere. However, this changes the distance between the grating vectors and the focal point of the spherical reference beam. In order to satisfy the imaging condition of the stored gratings, we need to move the focal point of the spherical reference until the Bragg-matched gratings have the correct displacement from the focal point. Just like before, only a portion of the entire hologram can be Bragg-matched at a time and the reconstructed portion deflects at a slightly different angle when compared to the original signal beam.

Instead of trying to re-Bragg-match the shrunken gratings during reconstruction, is there a way to prevent the gratings from shrinking in the first place? Specifically, can we use the ray nature of the spherical reference beam to record shrinkage insensitive hologram? In Chapter 3.2.3 we stated that for gratings recorded with symmetric rays (with respect to the recording material's surface normal), shrinkage has no effect on its Bragg condition. Figure 3.33 (a) shows a planewave reference beam interfering with a high bandwidth signal beam to produce a high density hologram (neglecting Snell's law). The angle of the reference beam is symmetric with the center ray of the signal beam.

Therefore, that portion of the hologram will remain Bragg matched after shrinkage. This

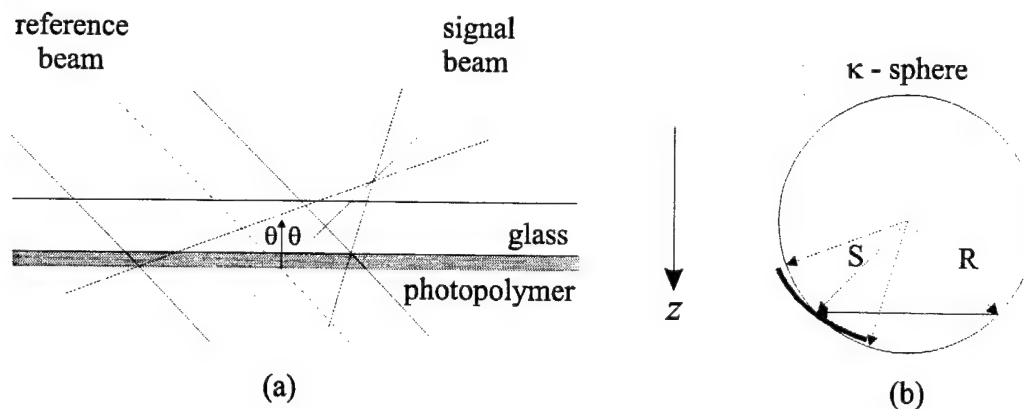is shown in the k-sphere diagram of Figure 3.33 (b).

Figure 3.33 : A planewave reference beam recording a hologram with a high bandwidth signal beam. (a) The recording geometry. (b) The κ-sphere diagram after shrinkage.

However, the planewave reference beam has only one incident angle so it cannot

record symmetric holograms with all the rays in the signal beam. This is where the

spherical reference beam comes in. Figure 3.34 (a) shows a spherical reference beam

recording a hologram with a high bandwidth signal beam. The focal point of the signal

beam is in front of the recording material and the focal point of the spherical reference

beam is behind. The lenses that produce the cone of rays for the signal and reference

arms have the same $f/\#$. When the rays are positioned properly as shown, the entire

hologram is formed by symmetric rays. More specifically, a pair of symmetric signal and

reference rays form a single grating at a localized area. This is repeated through out the

entire hologram. Since the gratings have no z-components, photopolymer shrinkage has

no effect on their Bragg conditions. This is shown in the κ-sphere diagram of Figure 3.34
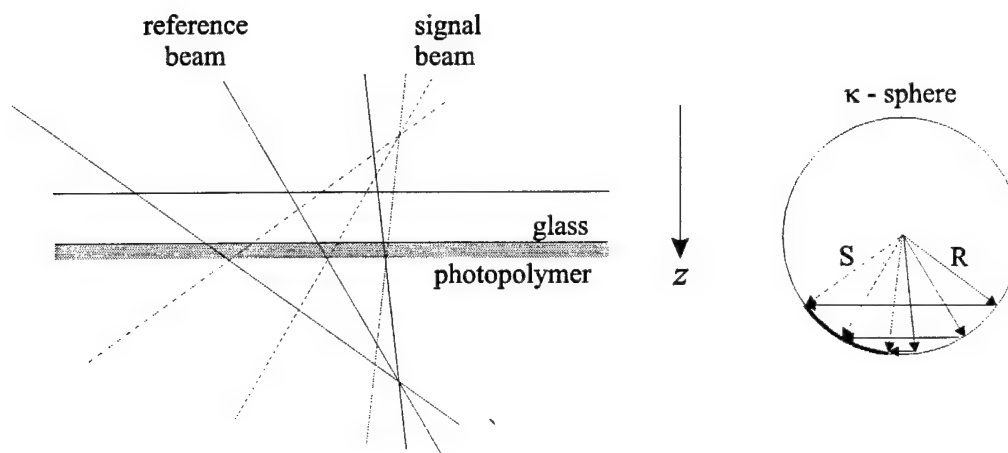
(b) for three grating vectors.

Figure 3.34 : A spherical reference beam recording a hologram with a high bandwidth signal beam in the shrinkage insensitive geometry. (a) The rays of the signal and reference beams in the setup. (b) The κ-sphere diagram after shrinkage.

To demonstrate that this shrinkage insensitive geometry actually works, we used basically the same high density setup as shown in Chapter 3.2. A piece of photopolymer is placed slightly after the focal point of the signal beam. To show the Bragg-mismatch effect induced by photopolymer shrinkage, we first use a planewave to write a hologram with the high bandwidth signal beam (Figure 3.35 (a)). The center ray of the signal beam and the reference planewave have an incident angle of $\pm 30°$ (outside angle) with respect to the surface normal of the photopolymer. The reconstructed hologram is viewed by a CCD detector placed at the detector plane of the Nikon lenses. Figure 3.35 (b) shows the shrinkage insensitive recording geometry. Instead of using a planewave to record the hologram, a spherical reference beam is generated by a $f/1.4$ CCD camera lens. The rays are positioned so that the signal beam spot on the photopolymer is covered by the reference beam spot and the center ray of the signal and reference beams overlap symmetrically with respect to the photopolymer's surface normal (at $\pm 30°$).
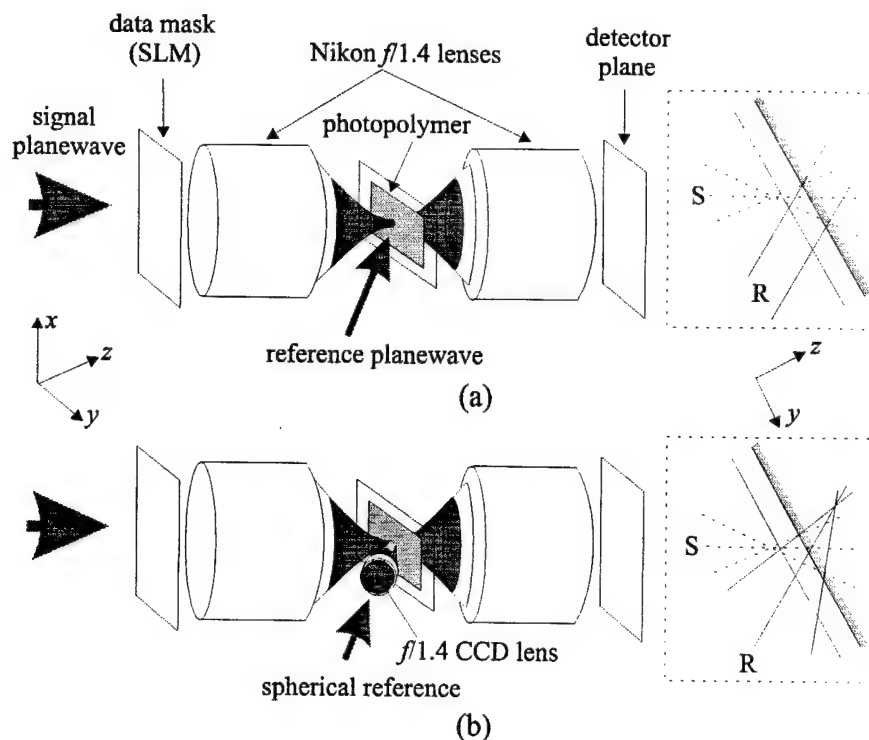
Figure 3.35 : Setups used to demonstrate the effects of photopolymer shrinkage. (a) High density setup using a planewave reference beam. (b) High density setup in the shrinkage insensitive geometry.

Figure 3.36 shows the resulting reconstructions and an image of the original data mask for comparison. Since the CCD array we used is much smaller than the entire aperture of the Nikon camera lens, we physically moved the detector array across the aperture to get the images shown in Figure 3.36. The 'original image' is the data mask imaged through the system without any recording material. The second image is the reconstruction of a hologram recorded with a planewave reference beam, using the setup shown in Figure 3.35 (a). Nearly half of the reconstruction is dark due to the shrinkage induced Bragg-mismatch problem. The dark half can be re-Bragg-matched by slightly changing the reference beam angle in the y-z plane, but then the other half would go dark. The last image shown in Figure 3.36 is the reconstruction of a hologram recorded with a spherical reference beam in the shrinkage insensitive geometry (Figure 3.35 (b)). The

entire reconstruction is visible at the same time so there is no need for any re-Bragg-matching. Furthermore, the reconstructed image is not distorted or compressed in any direction since the gratings did not suffer any shrinkage effect. This shows that the shrinkage insensitive geometry works and we will apply it to record shift multiplexed holograms in the photopolymer based holographic 3-D disks later.

The shrinkage insensitive setup shown in Figure 3.35 (b) has the signal beam focused in front of the recording material and the reference beam focused behind. Of course the same principles would still work the other way around. We choose to focus the signal beam in front of the recording material because that usually yields the best signal beam uniformity. For the reference beam, we purposely put a one-sided Gaussian profile on the planewave before the reference lens to smooth its intensity on the photopolymer.

replace with figure3-36.ps

We can compute the theoretical diffracted power as a function of hologram shift for the shrinkage insensitive geometry. For shifting parallel to the interaction plane, this is done by summing the diffracted power of each grating according to the angular selectivity function. The reconstruction intensity of a grating as a function of angular change in the reference ray is defined as:

$$I(\Delta\theta) = \left(\text{sinc}\left(\frac{L}{\lambda} \frac{\sin(\theta_r + \theta_s)}{\cos\theta_s} \Delta\theta \right) \right)^2$$ (3.16)

where $\text{sinc}(x) = \sin(\pi x)/\pi x$, L is the thickness of the recording material, λ is the wavelength, θ_r (θ_s) is the angle of the reference ray (signal ray) with respect to the surface normal of the recording material, and $\Delta\theta$ is the angular change in the reference ray as the recording material is shifted. Since the signal and reference rays are symmetric, Eq. 3.16 cane be further simplified to:

$$I(\Delta\theta) = \left(\text{sinc}\left(\frac{L}{\lambda} \frac{\sin(2\theta')}{\cos\theta'} \Delta\theta \right) \right)^2$$ (3.17)

where θ' is the angle between either the signal or the reference ray and the surface normal of the recording material. Now we need to determine $\Delta\theta$, the angular change in the reference ray as a function of recording material shift. After recording a hologram using the shift insensitive geometry, the same reference beam is brought back to reconstruct the hologram as shown in Figure 3.37. Snell's refraction is not shown in Figure 3.37 to simplify the diagram (the three rays shown do not actually intersect at the same location if refraction is taken into account). In Figure 3.37, θ is the angle of the largest ray measured with respect to the surface normal of the recording material, ϕ is the angle of the smallest

ray measured with respect to the surface plane of the recording material, and y is the spot

size of the spherical reference beam on the recording material.
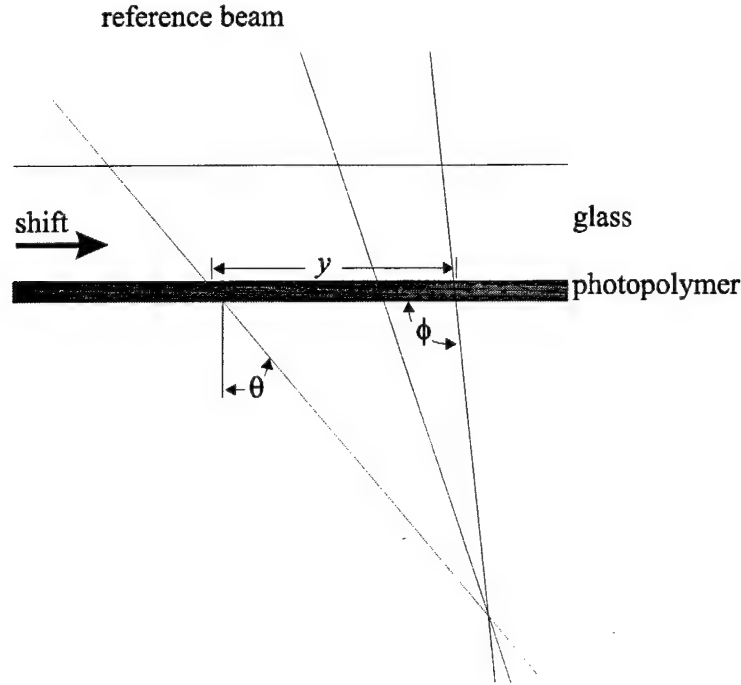
reference beam

shift

glass

y

photopolymer

φ

θ

Figure 3.37 : The angular change in the reference ray experienced by the shifted gratings for a spherical reference beam incident as shown.

Using simple trigonometry, the angular change for the grating recorded by a

reference ray with an angle θ' with respect to the surface normal of the recording

material, as the recording material is shifted from left to right is:

$$\Delta\theta \cong \frac{\cos\theta'}{\sin\phi}\frac{\cos(\theta-\phi)}{\cos\theta}\frac{\Delta y}{y}$$

(3.18)

where Δy is the shift distance. Substituting Eq. 3.18 into Eq. 3.17, we get:

$$I(\Delta y) = \left(\text{sinc}\left(\frac{L}{\lambda}\frac{\sin(2\theta')}{\sin\phi}\frac{\cos(\theta-\phi)}{\cos\theta}\frac{\Delta y}{y}\right)\right)^2$$

(3.19)

Eq. 3.19 is the theoretical diffracted power as a function of recording material

shift, for the grating recorded by the reference ray with an angle of θ' with respect to the

surface normal of the recording material. In order to get a complete picture of what is going on with the entire hologram as the recording material is shifted, we need to integrate the diffracted powers from all the gratings as a function of material shift. For computational purposes, this integration is usually broken down to a sum over a finite number of reference ray angles, such as

$$I(\Delta y) = \sum_{\theta'} \left(\text{sinc}\left(\frac{L}{\lambda} \frac{\sin(2\theta')}{\sin\phi} \frac{\cos(\theta - \phi)}{\cos\theta} \frac{\Delta y}{y} \right) \right)^2 \qquad (3.20)$$

where θ' is in the range of $90°-\phi$ to θ.

Figure 3.38 shows the diffracted power as a function of hologram shift (in the Bragg direction) for a high density hologram recorded using the same setup as Figure 3.35 (b). For this experiment, the $f/1.4$ CCD camera lens was moved closer to the DuPont HRF-150-100 photopolymer to project a larger reference beam spot size than the signal beam. This was done to ensure the signal beam is completely covered by the reference beam during recording (it will become more important later for the holographic 3-D disks with tilts and wobbles). The diameter of the reference spot on the photopolymer was ~ 4 mm while the diameter of the signal spot was 1.5 mm. For comparison, the theoretical diffracted power as a function of hologram shift for a hologram recorded with a reference spot size of $y = 1.5$ mm, $\theta = 30.1°$ and $\phi = 96.5°$ (inside angles) is also plotted in Figure 3.38. By moving the reference beam lens closer to the photopolymer to get a bigger spot size, the focal point of the spherical reference beam becomes displaced further away from the recording material. This increases the amount of material shift required to obtain the appropriate $\Delta\theta$ for Bragg-mismatching.

Therefore, the experimental selectivity curve is much wider than the theoretical prediction.
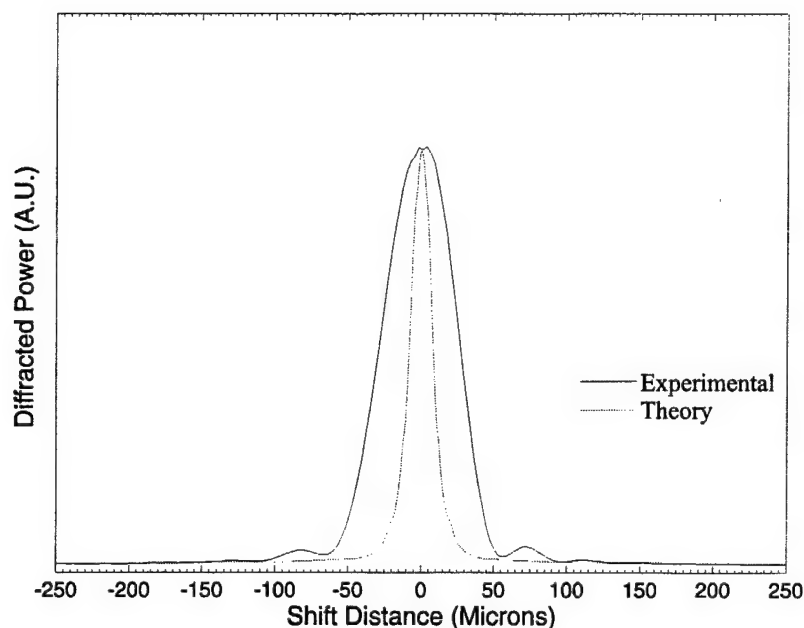

Figure 3.38 : Diffracted power as a function of shift distance for a high density hologram recorded with a $f/1.4$ reference beam lens.

Not only is the theoretical selectivity curve narrower than the experimental, it also doesn't have any side lobes. This is because the nulls of the sinc function for each grating does not occur at the same shift distance. Therefore, when the diffracted powers are summed together, the nulls get filled. Eq. 3.20 assumes each ray in the reference beam records similar strength holograms. In reality, since the reference beam is focused, the gratings near the center of the hologram are usually stronger. This in effect narrows the spectrum of reference rays that make real contributions to the diffracted power. When the strength of each grating was weighted by a Gaussian profile in Eq. 3.20, the theoretical diffracted power curve also had side lobes.

Notice the experimental selectivity curve of Figure 3.38 is asymmetric with respect to zero shift distance. The first null occurs at 57 µm for positive shift and -65 µm

for negative shift. This is because the hologram wasn't recorded in the shrinkage insensitive geometry. Since the $f/1.4$ CCD camera lens for the reference beam was moved closer to the recording material, the reference rays were no longer exactly symmetric with the signal rays. Therefore, most gratings have some z-components which shrunk during recording.

In order to maintain a larger reference spot size on the photopolymer than the signal beam and record in the shrinkage insensitive geometry, we need a lower $f/\#$ lens in the reference arm. Figure 3.39 shows that a lower $f/\#$ lens in the reference arm allows us to record holograms in the shrinkage insensitive geometry while illuminating a larger area of the photopolymer. We found a small (16.5 mm in length and 15.9 mm in diameter) projection lens from Universe Kogaku Inc. (Part Number TK-11) with a $f/\#$ of 1.1 that could be used in the reference arm to produce a spherical reference beam. Figure 3.40 shows the resulting selectivity curve for a hologram recorded with the $f/1.1$ projection lens.
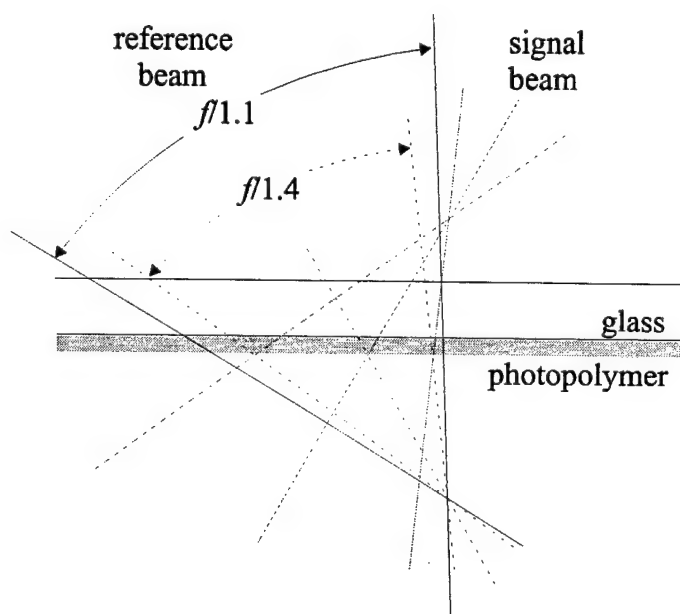
Figure 3.39 : Using a lower *f/#* lens in the reference arm (as compared to the signal arm) for shift multiplexing in the shrinkage insensitive geometry.
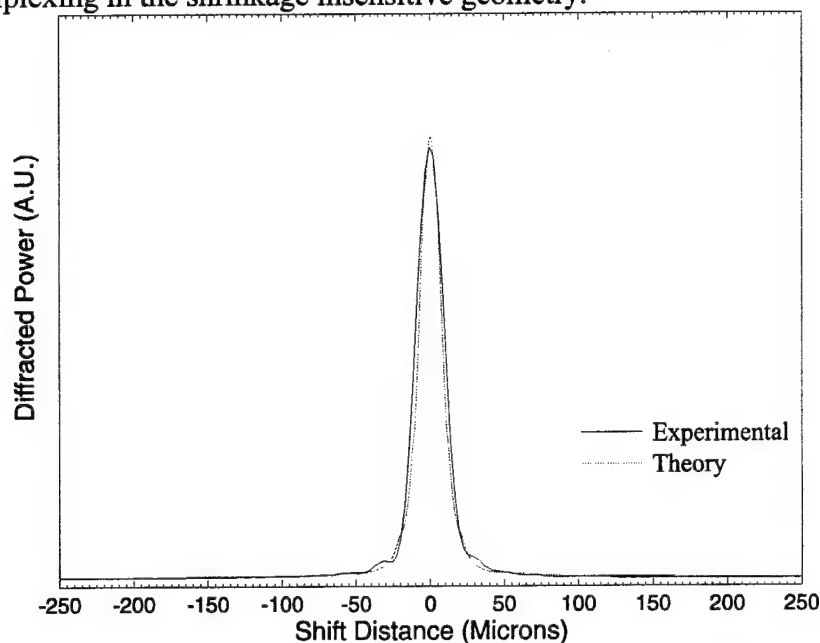

Figure 3.40 : Diffracted power as a function of shift distance for a high density hologram recorded with a *f/*1.1 reference beam lens.

Figure 3.40 shows that the experimental selectivity curve became much narrower and more symmetric with the *f/*1.1 projection lens in the reference arm. The reference beam diameter on the photopolymer was approximately 2 mm, just a little bit bigger than the theoretical diameter of 1.83 mm for the shrinkage insensitive geometry using a *f/*1.1 lens. Therefore, the theoretical selectivity curve is still slightly narrower than the experimental. According to the experimental selectivity curve in Figure 3.40, another hologram could be shift multiplexed, with little cross-talk noise, by shifting the recording material in the Bragg direction by 50 or more micron. This means that we could have a new hologram every 50 micron in shift, along the track of a holographic 3-D disk.

For the shrinkage insensitive geometry, the holograms are recorded with signal and reference beams of the same bandwidth. Therefore, the amount of shift required in the Bragg degenerate direction before another hologram could be stored is half the

hologram height, or .75 mm for this setup. Figure 3.41 (a) shows the κ-sphere diagram during recording with three representative grating vectors and Figure 3.41 (b) shows the κ-sphere diagram after shifting in the Bragg degenerate direction by half the hologram height. Grating vectors with some x-components quickly Bragg-mismatch when shifted in the Bragg degenerate direction, leaving a strip of grating vectors that have no x-components on the κ-sphere. The .75 mm in orthogonal shift is enough to move the grating vectors that have no x-components out of the field of view of the detector array.
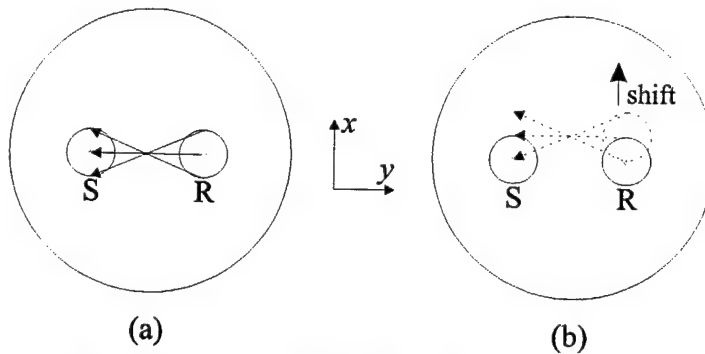

(a)   (b)

Figure 3.41 : κ-sphere representation of shift multiplexing using a spherical reference beam in the shrinkage insensitive geometry. (a) Recording a hologram between a signal and reference beam with the same bandwidth. (b) Shifting in the degenerate direction by half the hologram height during reconstruction.

3.3.4 Experimental Setup

After figuring out how shift multiplexing works, we are now ready to construct a photopolymer based, holographic 3-D disk recorder/player system using shift multiplexing. Figure 3.42 shows the schematic diagram of a shift multiplexed holographic 3-D disk setup. It is very similar to the setup used in the previous experiment and many of the same components are used. An E-beam lithographed chrome plate is used as the input spatial-light-modulator (SLM). The center-to-center spacing of the pixels is 45 μm, and the fill factor is 100%. Nikon f/1.4, 3.9-cm-aperature camera lenses

are used for imaging (the chrome plate is pressed up against the Nikon lens to ensure that all the pixels within the lens aperture are captured). A total of 590,000 pixels fit in the apertures of the two Nikon lenses and a sharp image of the entire field is obtained at the detector plane. The recording material, DuPont's HRF-150 100 micron thick photopolymer, is laminated on the non anti-reflection coated side of a 1.2 mm thick transparent glass disk. The glass disks used in this experiment has the same physical dimensions as the compact disc. Figure 3.43 shows a picture of a photopolymer based holographic 3-D disk next to a cube of $LiNbO_3$. The glass disk laminated with the DuPont photopolymer is then attached to a microstepping-stepper motor. This particular stepper motor can be set to 125,000 steps per full revolution. This corresponds to a shift distance of approximately 3 microns per step at the outside radius of the disk. The entire glass disk and stepper motor assembly is mounted on a translation stage so the disk could be moved up or down to access different tracks on the disk. The spherical reference beam is formed by using a $f/1.1$ miniature projection lens. The holograms are recorded with the recording material slightly past the Fourier transform plane of the Nikon lenses. At that position, the diameter of the signal beam is 1.5 mm and its spatial uniformity is much better than at the exact Fourier plane. The spherical reference beam generated by the $f/1.1$ projection lens focuses behind the recording material in the shrinkage insensitive geometry discussed in Chapter 3.3.3.3. The center rays of the signal and reference beams each makes an angle of 30 degrees (outside) with respect to the disk's surface normal. The intensity of the signal and reference beams are 10 $\mu W/cm^2$ and 80 $\mu W/cm^2$ respectively, measured from the collimated planewaves before they become focused by

the Nikon and projection lenses. At the detector plane, a regular video rate CCD camera

is used to read the reconstructed holograms. Figure 3.44 shows a picture of the setup.
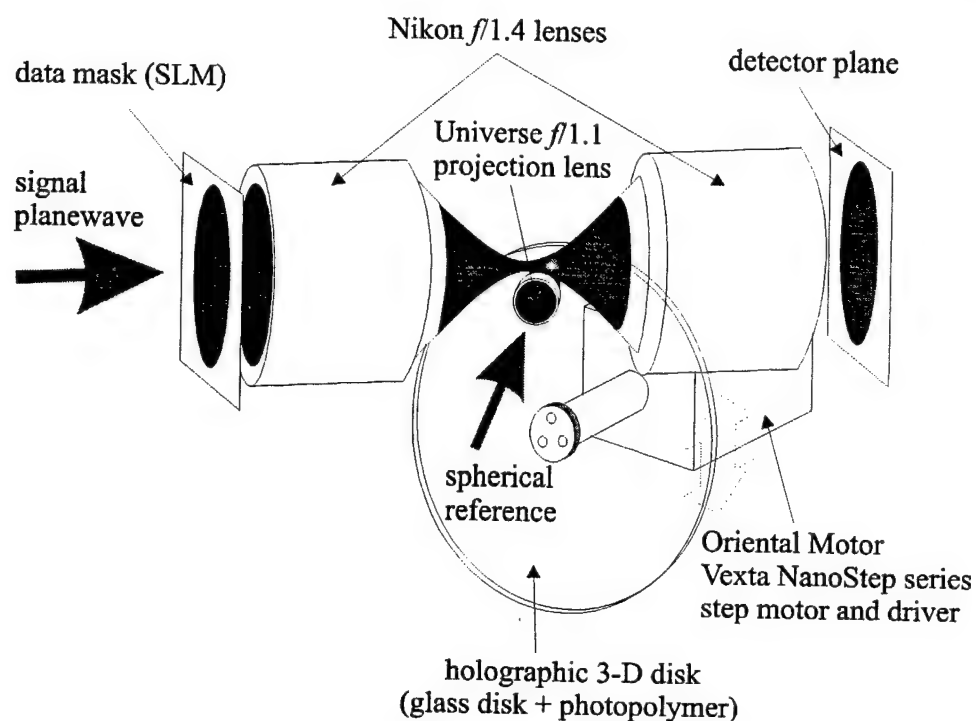

Figure 3.42 : Schematic diagram of a holographic 3-D disk setup using shift multiplexing.
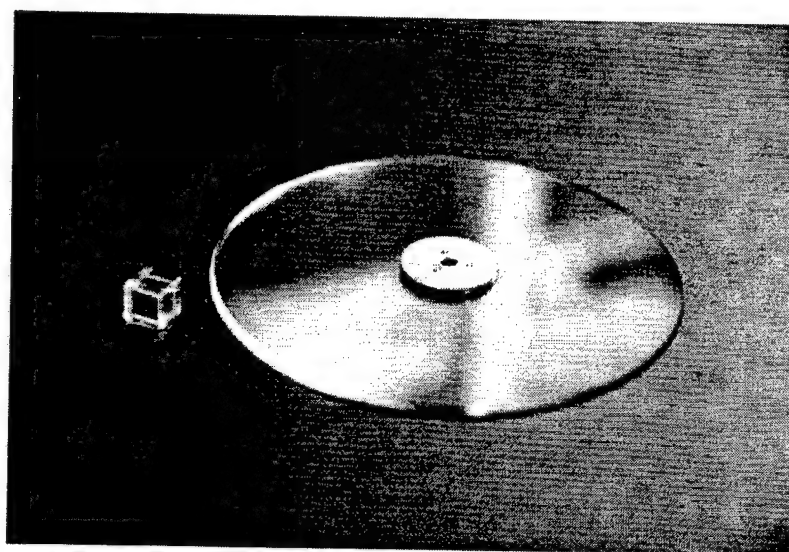

Figure 3.43 : A picture of a photopolymer based holographic 3-D disk next to a cube of LiNbO$_3$.
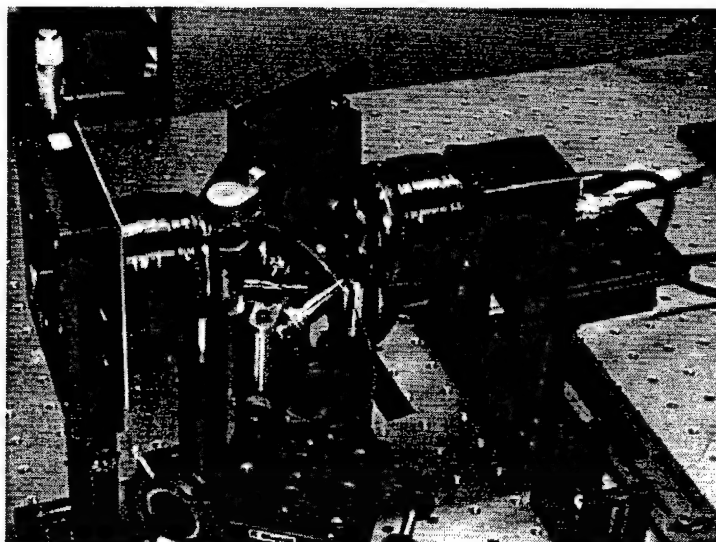
Figure 3.44 : A picture of the holographic 3-D disk setup using shift multiplexing.

We know from Chapter 3.2 that the 100 micron thick DuPont photopolymer can achieve a surface density of 10 bits/μm^2 with sufficient SNR. Therefore, we will tune the new setup to achieve it again. For the holographic 3-D disk, the in-track density is defined as the number of pixels in a hologram over the effective area of the hologram. Since the holograms are recorded in partially overlapping areas by using shift multiplexing (see Figure 3.45), the effective area can be computed as the height of the hologram times the shift distance required to separate the next hologram. It is similar to having a distinct hologram stored within each shift distance. To achieve the goal of 10 bits/μm^2 with the new setup, we will first start with an in-track surface density of 5 bits/μm^2. Then by overlapping the tracks, boost the surface density to the target 10 bits/μm^2. For this setup, we have again 590,000 pixels in each hologram over a hologram height of 1.5 mm. Therefore, to achieve an in-track surface density of 5 bits/μm^2, a maximum shift distance of 78 micron between holograms is allowed. To double the surface density to 10 bits/μm^2, we overlap the tracks by storing one track at half the

hologram height above and another track at half the hologram height below, with respect to the middle track. This in effect give us two completely overlapping tracks for a surface density of 10 bits/μm^2.
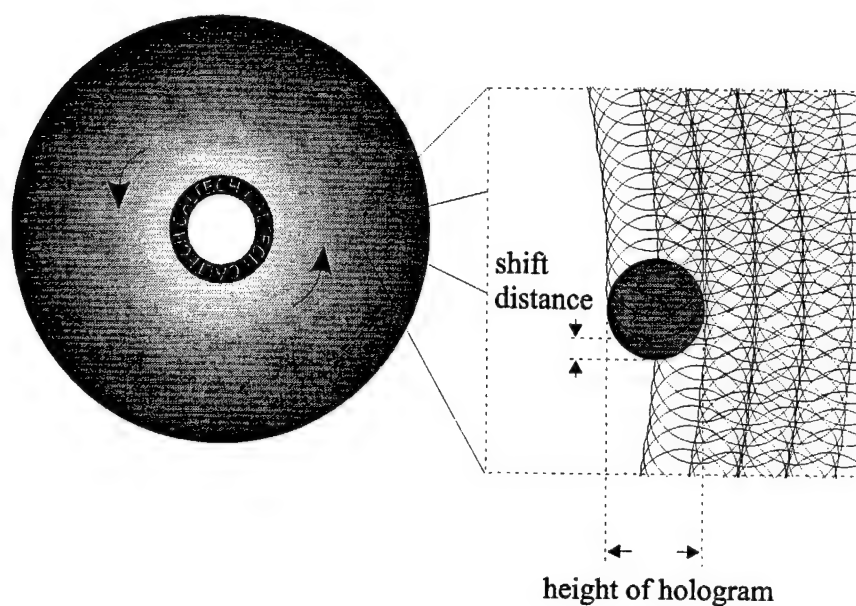

Figure 3.45 : The layout of overlapping holograms in tracks and overlapping tracks in the holographic 3-D disk.

3.3.5 Experimental results

So far, a surface density of 10 bits/μm^2 has been achieved on paper by determining the shift distances required. From Figure 3.40 of Chapter 3.3.3.3, we know it is possible to pack in-track holograms as closely as 50 μm in separation and overlap additional tracks at half the hologram height above and below the current track. To demonstrate that we can actually store a hologram every 78 μm in in-track shift without too much cross-talk with the setup shown in Figure 3.44, we recorded one hologram and measured its diffraction efficiency as a function of shift distance (Figure 3.46). In Figure

3.46, the 78 µm in-track shift between holograms actually corresponds to approximately the fourth null. Therefore very little cross-talk is expected (the next hologram would be recorded at position 688 and 532 micron). For track overlapping, we observed that by shifting the recording material radially by .75 mm (half the hologram height), is enough to deflect the reconstruction of the Bragg-matched strip completely off of the detector plane. Therefore, another track could be recorded at half the hologram height above and below the current track with essentially zero cross-talk.
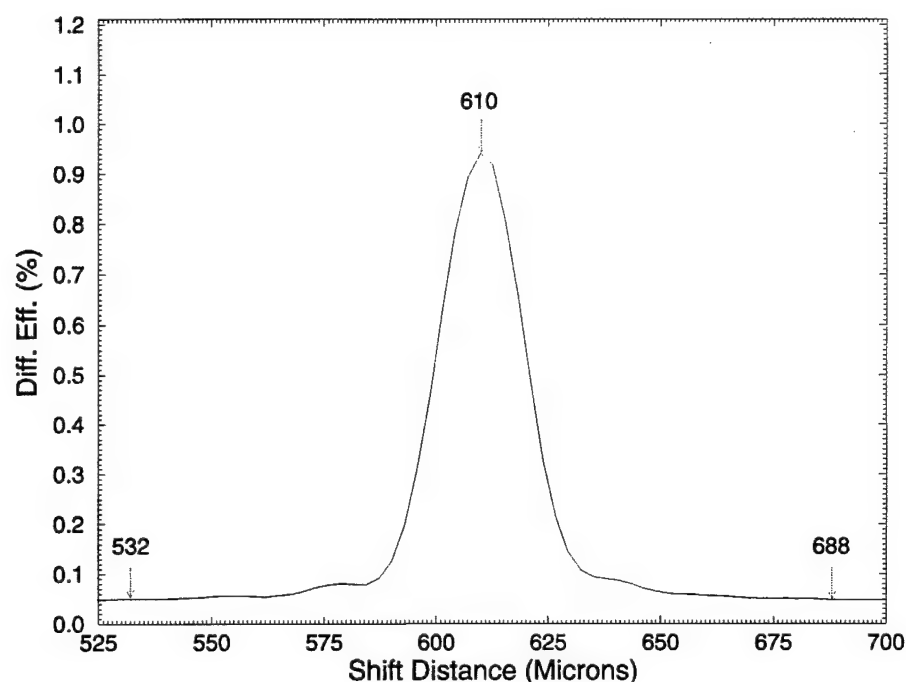

Figure 3.46 : Diffracted power as a function of shift distance for a high density hologram recorded with our shift multiplexed holographic 3-D disk setup.

To experimentally demonstrate 10 bits/µm², three overlapping tracks separated radially by .75 mm were stored. In each track, 100 holograms were recorded with an in-track separation of 78 µm (some initial holograms in the middle track were omitted so cross-talk noise could be characterized). Before recording, the photopolymer disk was uniformly sensitized in room light for 5 minutes. The holograms were then recorded with

a constant 500 ms exposure per hologram and the middle track was recorded first, then the upper and lower. The average diffraction efficiency of the holograms was approximately half of a percent. Figure 3.47 shows the comb function of the stored holograms (diffracted power as a function of shift distance). The strength of the holograms were very uniform by using the simple constant exposure schedule for the first track recorded (middle track). The holograms recorded later in the upper and lower tracks were slightly weaker due to the limited dynamic range of the material. To boost the strength of these holograms, they could have been recorded for longer durations.

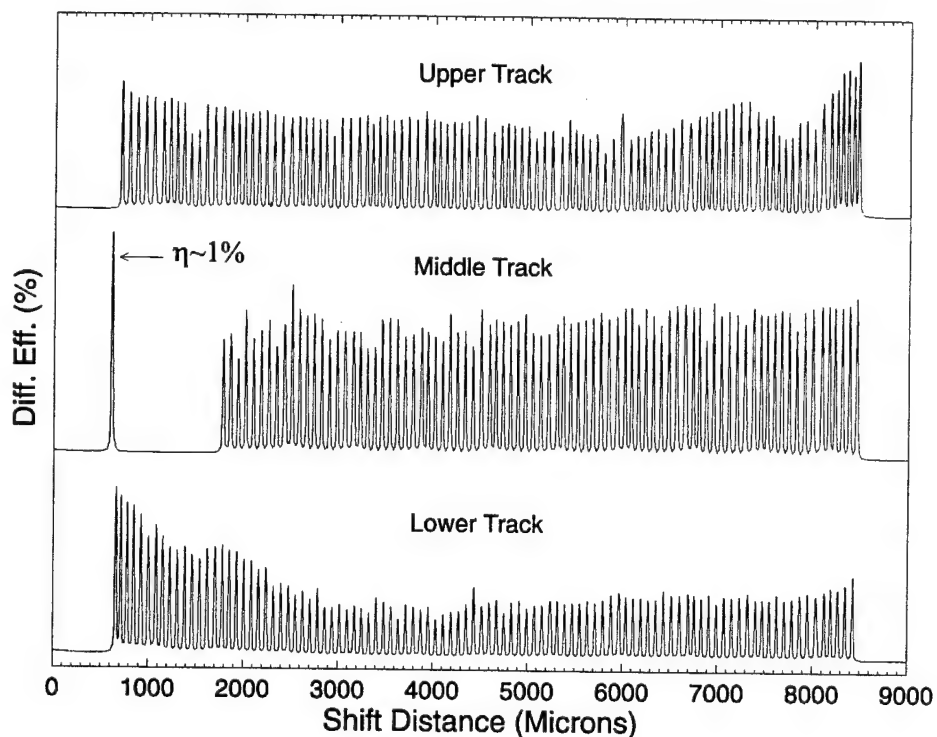Figure 3.47 : Diffracted power as a function of shift distance for three overlapping track, each track containing approximately 100 holograms.

One of the reasons why the diffraction efficiency of the upper and lower tracks are less than the middle track is because of the "run-time" effect discussed in Chapter 2.4.2. From pre-exposure to recording to UV curing, the room light is kept on to maintain the

sensitivity of the photopolymer. The amount of time spent on recording 100 holograms in a single track is around 100 seconds (500 ms exposure time + 500 ms delay to allow the stepper motor to stabilize for each hologram). This means by the time the lower track is started, the photopolymer has already been in the room light for about 8 minutes. During this time, the photosensitizing dye is slowing being depleted by the room light, causing the photopolymer to be less sensitive. We also tried other sensitizing techniques such as letting the larger reference spot pre-expose the next hologram area automatically during the recording process. This technique pre-exposes the minimum amount of photopolymer necessary for hologram formation. With the right combination of exposure time and incident intensity, we were able to get this method to work but not always consistently from disk to disk.

Figure 3.48 shows a reconstructed hologram from the middle of the middle track, where the cross-talk noises are most pronounce. Since the CCD array used was much smaller than the entire reconstruction, only a window of the reconstruction can be read out at a time. The three frames in Figure 3.48 correspond to the left, center and right sides of the reconstructed hologram. Visually, the quality of the reconstructed hologram is quite high and the reconstruction from the other holograms were all similar or better than this. To estimate the bit-error-rate of our system, we generated a histogram from the reconstructed hologram, using the same method as shown in Figure 3.10. The black area in Figure 3.49 represents the occurrences of the 'off' pixels while the shaded area represents the occurrences of the 'on' pixels, plotted against their intensity value. The two lobes are distinct, meaning that an intensity threshold could be selected to classify the

pixels as originally 'on' or 'off' without errors. The estimated probability of error derived from fitting the 'on' pixel histogram to a first order χ^2 distribution is $2.4\text{x}10^{-5}$.
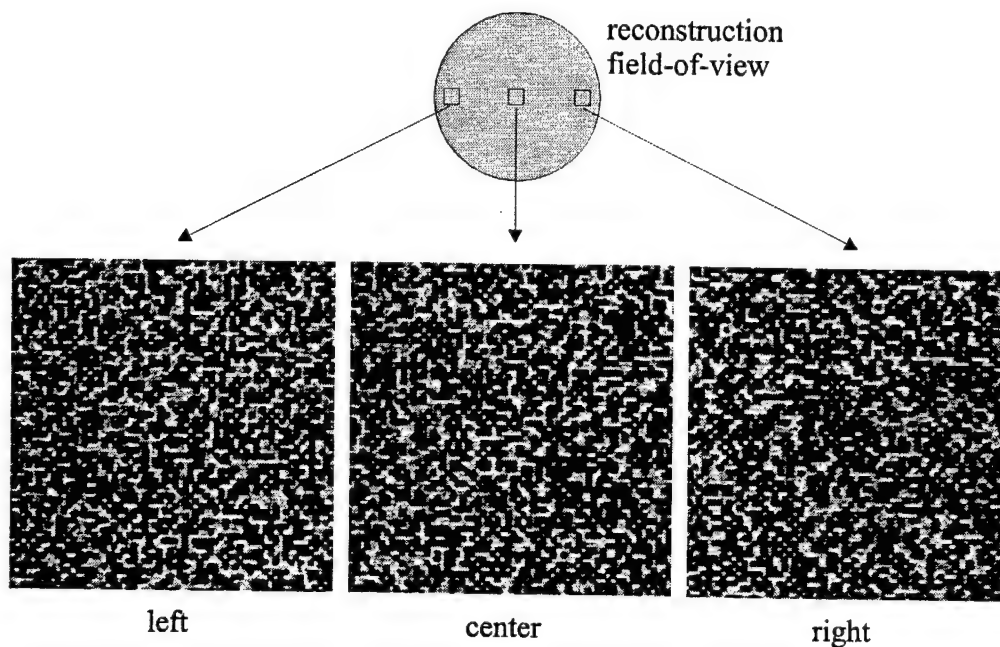

Figure 3.48 : The left, center, and right sides of a reconstructed hologram from the 10 bits/μm^2, photopolymer based holographic 3-D disk experiment.
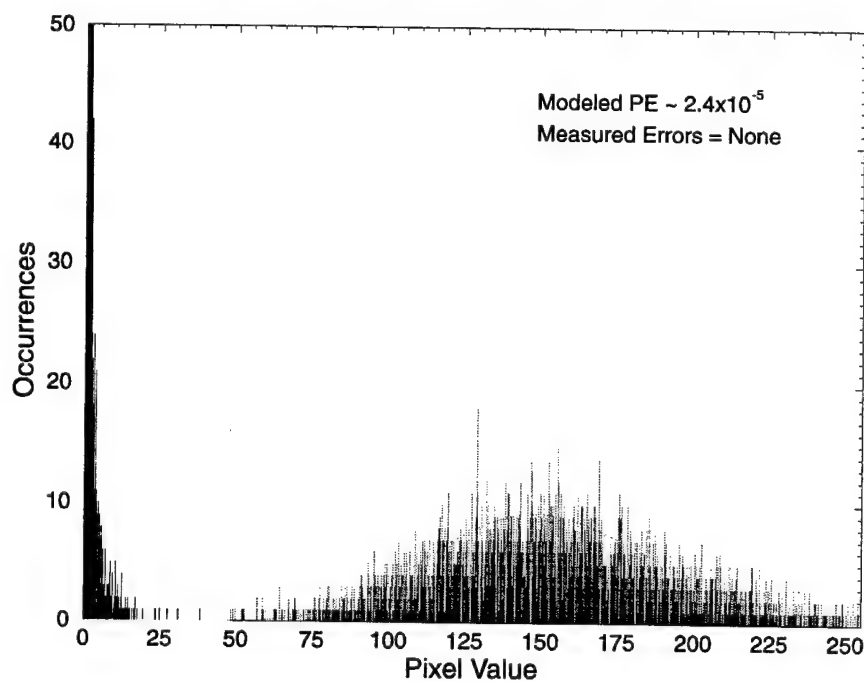

Figure 3.49 : Combined histogram obtained from three different sampled windows for the 10 bits/μm^2, photopolymer based holographic 3-D disk experiment.

We also measured the SNR of the system under various experimental conditions as shown in Figure 3.50. The SNR for both the shift multiplexed holographic 3-D disk experiment and the previous angle+peristrophic multiplexed experiment are plotted for comparison. The first SNR measurement of the shift multiplexing holographic 3-D disk setup is 11.5, which is due to system noises only (lens aberrations, SLM imperfections, detector noise, scattering and multiple reflections from lenses and other optical components, laser non-uniformity and fluctuations, and SLM-to-CCD pixel misalignment). We obtained this by transmitting the signal beam through the system without any recording material. An improvement of 1.5 in SNR when compared with the angle+peristrophic multiplexing setup is due to the E-beam lithographed, chrome plated data mask used in the holographic 3-D disk setup. The optically lithographed data mask used in the previous setup did not have the uniformity nor the contrast of the E-beam lithographed mask. By placing the one side AR-coated 1.2 mm thick glass disk at the recording plane, the SNR drops to 10.9. This slight drop in SNR is mostly due to aberration and internal reflections caused by the disk substrate. However, the new setup has an additional SNR improvement of 0.4 over the old setup (when the SNR increase of 1.5 have already been taken into account). This small increase in SNR is due to the anti-reflection coating on the glass disk. In the old setup, regular 1 mm thick glass slides were used to support the photopolymer. We suspect that the SNR would have been even better had the glass disk been AR coated on both sides. However, this was not done since the photopolymer is laminated on one side of the disk and would have ruined the AR effect. When we introduced a piece of UV-cured photopolymer laminated on the disk substrate

at the recording plane, the SNR dropped further to 8.8. The major cause of this drop in SNR is most likely due to internal reflections within the film and the glass substrate and scattering caused by the photopolymer film. Since the same type of photopolymer was used in both experiments, the drop in SNR caused by the UV-cured photopolymer is the same for both setups.
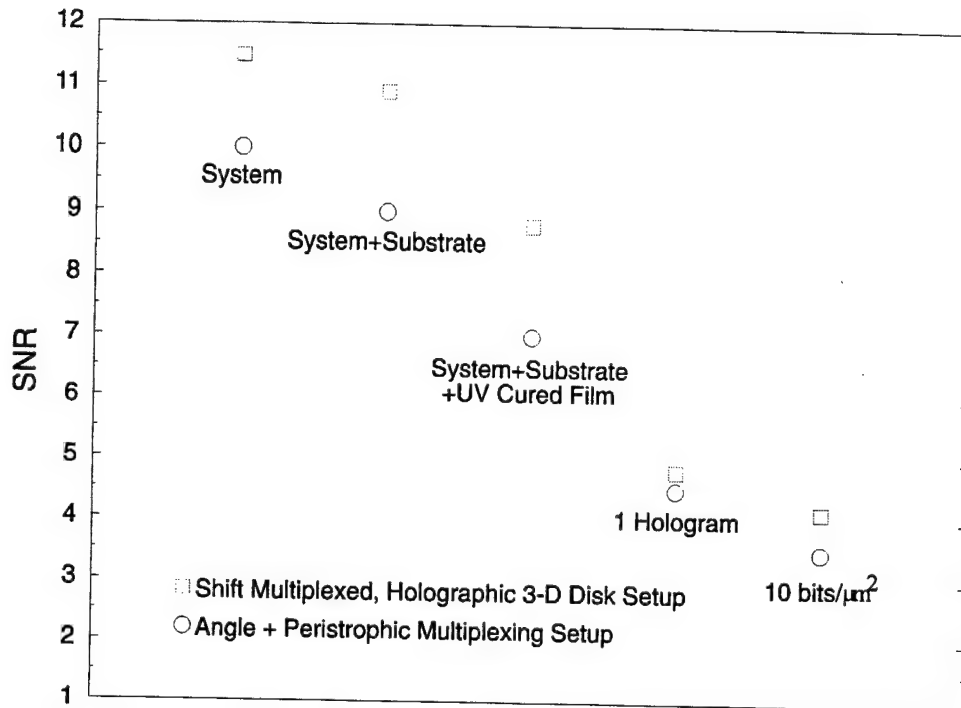
Figure 3.50 : SNR characterization for the shift multiplexed, photopolymer based holographic 3-D disk setup.

A SNR of approximately 4.8 was obtained when a single hologram was recorded on the disk and its reconstruction evaluated. No errors were observed in the reconstruction and the histogram of the 'on' and 'off' pixels were still clearly separated. However, when compared with the old setup, the lead in SNR enjoyed by the new setup has shrunk to only 0.3 (from a lead of 1.8 before). To a large extent, this significant drop in SNR (from 8.8 to 4.8) is due to internal reflections from the boundaries. Figure 3.51 shows a close-up diagram of the recording plane. The signal beam is a high bandwidth

(*f*/1.4) spherical beam that focuses in front of the recording material and the reference beam is an even higher bandwidth (*f*/1.1) spherical beam that focuses in the back. Unlike the angle+peristrophic multiplexing setup where the reference beam was just a planewave and internal reflections caused an interference pattern that was distributed across the entire hologram. The reflected spherical reference beam for the holographic 3-D disk setup re-focuses near the right side of the hologram. Therefore, the quality of a hologram stored in the geometry shown in Figure 3.51 will have poorer SNR on the right side when compared to the center or the left. For the one hologram experiment, the right side had an SNR of 4.2 while the left and the center had an SNR of 4.6 and 5.6, respectively, for an average SNR of 4.8. If there was a way to AR coat the photopolymer to reduce the back reflections, the SNR should improve significantly.

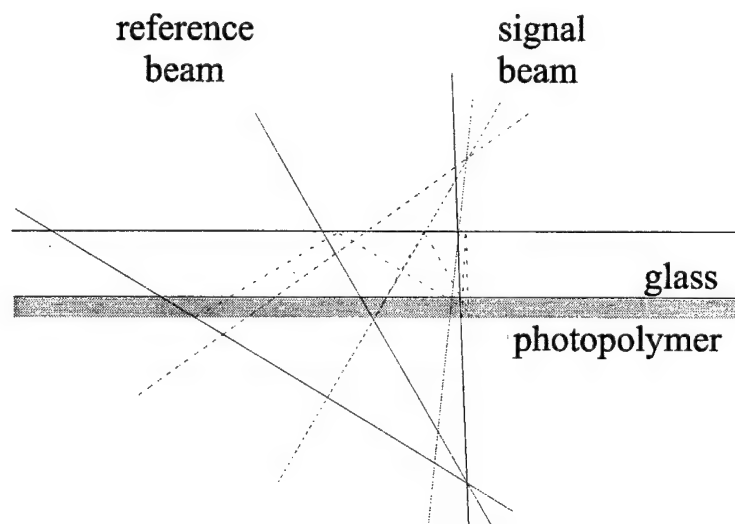
Figure 3.51 : Internal reflections within the holographic 3-D disk cause the right side of the stored hologram to have lower SNR.

For the final 10 bits/μm^2 surface density experiment with three overlapping tracks (100 holograms in each track), the measured SNR is 4.16. This is about 0.6 higher than the angle+peristrophic setup for the same surface density. Instead of maintaining a lead

of 0.3 in SNR for one hologram, the SNR actually increased by another 0.3 for the 10 bits/μm^2 experiment using shift multiplexing. This is probably because for the shift multiplexed holographic 3-D disk setup, less cross-talk noise is introduced between the in-track holograms when compared to the angle holograms of the angle+peristrophic setup. Furthermore, since the holograms are recorded in partially overlapping areas (instead of recording all the holograms at the same spot), the index modulation effect is probably less. The better SNR for the 10 bits/μm^2 shift multiplexed holographic 3-D disk experiment translates to a lower estimate bit-error-rate of 2.4×10^{-5}, more than 4 times lower than the angle+peristrophic experiment shown in Chapter 3.2.

3.4 100 bits/μm^2 High Density Experiment

Using Shift Multiplexing and a 1 mm Thick LiNbO$_3$

3.4.1 Introduction

In Chapter 3.3, a surface density of 10 bits/μm^2 was achieved with a shift multiplexed holographic 3-D disk system, using a 100 micron thick photopolymer. The storage capacity of a 120 mm photopolymer based holographic 3-D disk would be around 10 Gbytes (at 10 bits/μm^2). That is about 10 times higher than the capacity of the current compact disk. However, the next generation compact disc due in early 1997, the Digital-Versatile-Disc (DVD), will have a capacity of ~6 Gbytes per layer, for a maximum planned capacity of ~20 Gbytes per double-sided disk with two layers on each side. That translates to a surface density of approximately 20 bits/μm^2 for the 4 layered DVD disk. Unless we can get a much higher surface density than 20 bits/μm^2 with the holographic 3-

D disk system, there is no need to compete with the DVD for the same market with an inferior product.

So far the recording material we have used is only 100 micron in thickness. According to Eq. 3.15 and 3.20, the number of holograms you can pack in the same area is directly proportional to the thickness of the recording material. The thicker the recording material, the more holograms you can store in the same area and therefore, the higher the surface density. For example, if we were to use a 1 mm thick photopolymer based holographic 3-D disk, the surface density should increase by a factor of 10X to 100 bits/μm^2. Certainly a comfortable margin over the projected surface density of the DVD system. However, no 1 mm thick photopolymer exists at the current time so we will have to use another recording material as a substitute for now. A candidate that readily presents itself is LiNbO$_3$. Its photorefractive properties are well known and a 1 mm thick sample would be easy to obtain. LiNbO$_3$ can be polished to optical quality and AR coated on both sides to reduce reflections. However, some of its less desirable properties will force us to change our setup slightly. Due to the expense of acquiring large samples of LiNbO$_3$, we will not be able to make a LiNbO$_3$ based, holographic 3-D disk in the style shown in Figure 3.43. Instead, we will use a 2 cm by 2 cm piece of LiNbO$_3$ to demonstrate the system. For transmission geometry LiNbO$_3$, there is a preferred crystal axis for hologram formation. If the crystal is rotated like a disk during recording, the preferred axis will introduce a sinusoidal variation in the diffraction efficiency of the stored holograms [16]. Therefore we will not rotate the LiNbO$_3$ to achieve shift multiplexing. Instead, we will translate the crystal horizontally to record in-track holograms and vertically to overlap the tracks. Furthermore, since LiNbO$_3$ is not as

sensitive in the green region (532 nm) where we have done the previous experiments, we will move the setup to a large frame Argon laser that produces nearly 1 Watt of power at 488 nm.

The spirit of the experiment remains unchanged: to demonstrate that a surface density of 100 bits/μm^2 is possible with a 1 mm thick recording material by using shift multiplexing. The results obtained from this experiment can be applied to a photopolymer based, shift multiplexed holographic 3-D disk system when a 1 mm thick photopolymer becomes available.

3.4.2 Experimental Setup

The setup for this experiment is pretty much the same as the shift multiplexed holographic 3-D disk system shown in Chapter 3.3.4. A pair of Nikon $f/1.4$, 3.9-cm-aperture camera lenses were used to image an E-beam lithographed chrome plated data mask to a CCD detector. A total of 590,000 pixels fit in the apertures of the two Nikon lenses, and a sharp image of the entire field was obtained at the detector plane. The center-to-center spacing of the pixels was 45 μm and the fill factor was 100%. The recording material, 2 cm $\times$ 2 cm $\times$ 1 mm iron doped $LiNbO_3$ was mounted on two translation stages for shift multiplexing. Horizontal displacement (parallel to the plane of interaction) was provided by a Klinger CC1.2 linear stage and controller. The resolution of the Klinger linear stage is .1 μm/step and the scan time for 400 steps (40 μm) is about 8 seconds. Vertical displacement was provided by a manual translation stage. The spherical reference beam was formed by using a $f/1.1$ projection lens. The holograms were recorded with the $LiNbO_3$ slightly past the Fourier transform plane of the Nikon

lenses. At that position, the diameter of the signal beam was 1.5 mm on the entrance side and its spatial uniformity was much better than at the exact Fourier plane. Figure 3.52 shows a schematic diagram of the setup and Figure 3.53 shows a picture of the setup.
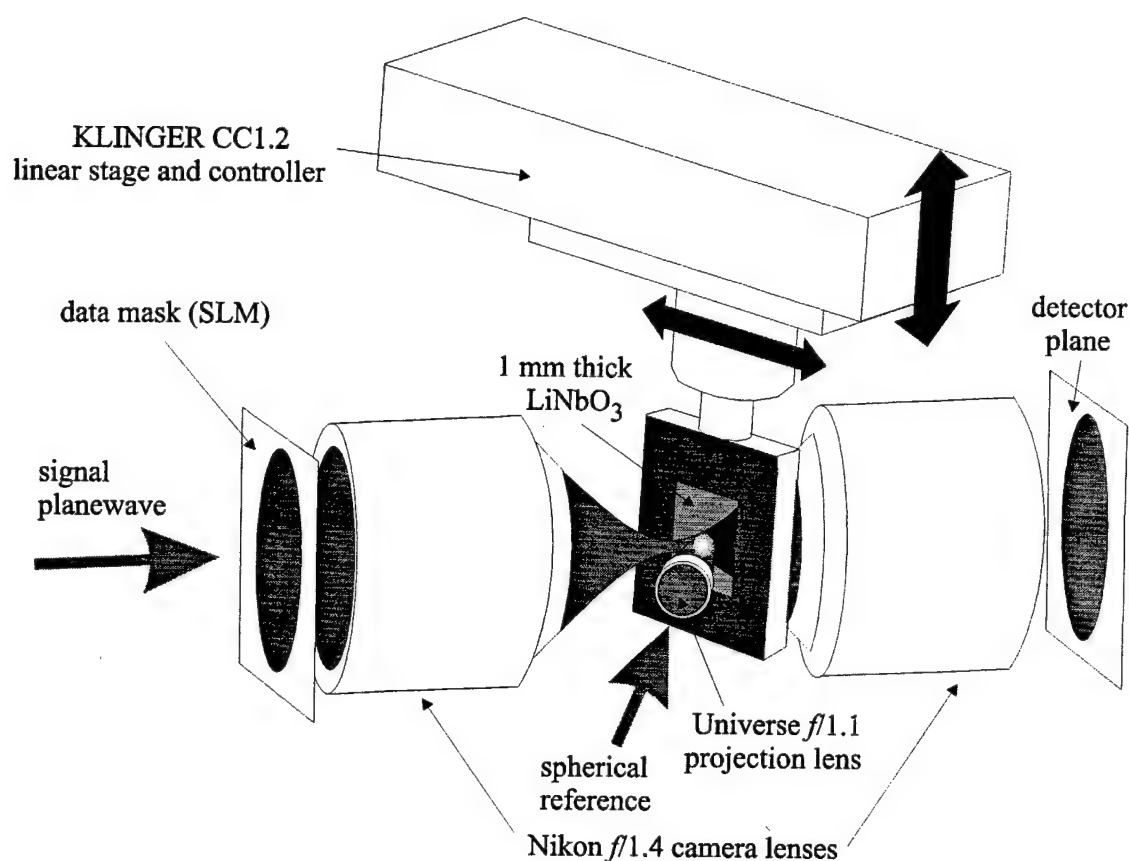

Figure 3.52 : Schematic diagram of a shift multiplexed high density setup using a 1 mm thick LiNbO$_3$.

One nice thing about LiNbO$_3$ is that it doesn't shrink when exposed to light. This frees us from the restrictive shrinkage insensitive geometry. For example, Figure 3.54 shows a recording geometry where both the signal and reference beams focus in front of the recording material. For this setup, the angle between the reference and signal rays for a localized grating remains relatively constant throughout the entire hologram. For the shrinkage insensitive geometry of Figure 3.34, the angle between the reference and signal rays is small on the right side and gets progressively bigger toward the left side of the

hologram.   Therefore, the Bragg selectivity of the entire hologram should be more uniform for the setup shown in Figure 3.54 when shifted parallel to the plane of interaction.   We tried this geometry with the center ray of the signal (reference) beam making an angle of 45° (35°) with respect to the recording material's surface normal.
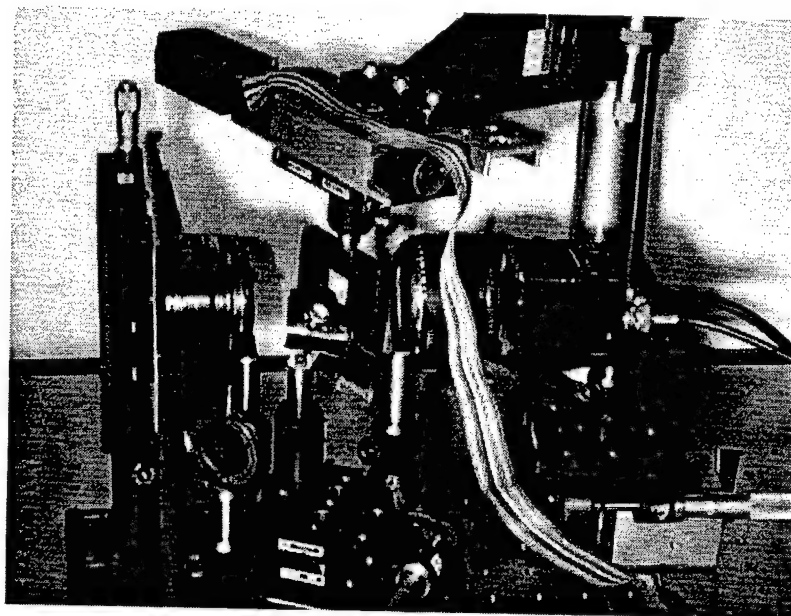

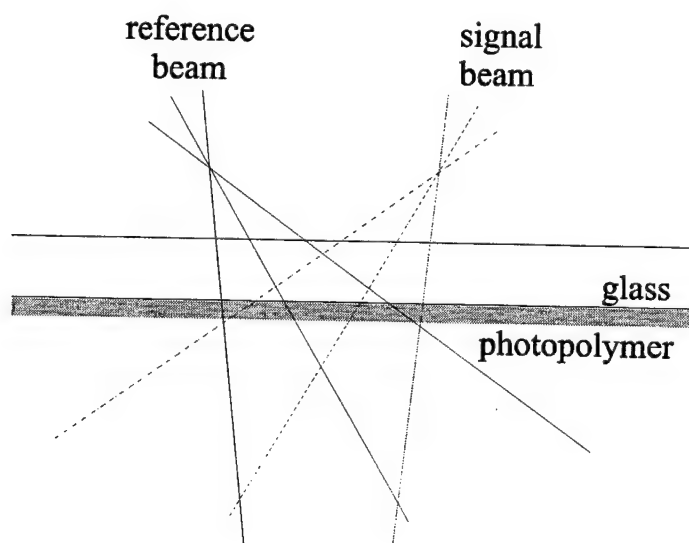Figure 3.53 : A picture of the shift multiplexed high density setup using a 1 mm thick LiNbO₃.

Figure 3.54 : A possible geometry for recording shift multiplexed holograms in LiNbO₃.

What we found was that the right side of the stored hologram Bragg-mismatched much more slowly than the left side, despite our effort to make the angle between the reference and signal rays more uniform across the entire hologram. Upon further analysis, it became clear why this happened. In order to have good angular selectivity (Eq. 3.7), it is desirable to have a large θ_s. The gratings on the right side of the hologram have a smaller θ_s when compared to the left side. Therefore, the angular selectivity for the gratings on the right side of the hologram is poorer than the left. Furthermore, since the focal spot of the spherical reference is in front of the recording material, it is further away from the right side of the hologram. This causes the rate of change in the reference ray angle to occur more slowly for the right side when the recording material is shifted. These two effects combine to make the right side Bragg-mismatch noticeably slower than the left side. For the shrinkage insensitive geometry of Figure 3.34, θ_s on the right side is also smaller than the left. However, the focal point of the spherical reference beam is closer to the right side when it is focused behind the recording material. Therefore, these two effects cancel each other to some degree, producing a more uniformly Bragg-mismatch hologram when shifted parallel to the plane of interaction. Hence, even though LiNbO$_3$ does not shrink, we will still use the shrinkage insensitive geometry for this setup by focusing the spherical reference beam behind the recording material.

The angular selectivity equation when written explicitly for wavelength and angles measured from outside of the recording material is :

$$\Delta\theta = \sin^{-1}\left(n\sin\left(\frac{\lambda}{nL}\frac{\cos\left(\sin^{-1}\left(\frac{\sin\theta_s}{n}\right)\right)}{\sin\left(\sin^{-1}\left(\frac{\sin\theta_s}{n}\right)+\sin^{-1}\left(\frac{\sin\theta_r}{n}\right)\right)}+\sin^{-1}\left(\frac{\sin\theta_r}{n}\right)\right)\right)-\theta_r \qquad (3.21)$$

where n is the index refraction of the recording material, λ is the wavelength of the laser, L is the thickness of the recording material, and θ_s (θ_r) is the angle between the signal (reference) ray and the surface normal of the recording material. For the photopolymer based, shift multiplexed holographic 3-D disk setup shown in Chapter 3.3, $n = 1.525$, $\lambda = 532$ nm, $L = 100$ μm, and the center ray of the signal and reference each made an angle of 30° with respect to the surface normal of the recording material. For the shift multiplexed, high density setup using LiNbO$_3$, the index of refraction of LiNbO$_3$ is 2.25 (2.35 for the ordinary axis), $\lambda = 488$ nm, and $L = 1$ mm. The shorter wavelength and higher index of refraction (compresses the wavelength inside the recording material) helps to sharpen the selectivity function. However, the higher index of refraction also decreases the angle of the signal and reference rays inside the recording material, which hurts the selectivity function. From Eq. 3.21, it is unclear if the overall effect of using a 488 nm laser and a recording material with an index of refraction of 2.25 increases or decreases the amount of $\Delta\theta$ required to reach the first null in Bragg condition. By using the parameters of the photopolymer based holographic 3-D disk, we get a $\Delta\theta$ of 0.5° from Eq. 3.21 for the grating recorded between the center rays. Using $\theta_r = \theta_s = 30^\circ$, $n = 2.25$, $L = 1$ mm, and $\lambda = 488$ nm, we get a $\Delta\theta$ of 0.071° for the LiNbO$_3$ high density setup. This means that the higher index of refraction of the LiNbO$_3$ crystal hurts the selectivity function more than it helps since the 1 mm thick recording material did not reduce $\Delta\theta$ by a factor of 10X. Perhaps we can regain some selectivity by using larger incident angles for the reference and signal beams. According to Eq. 3.21, the best $\Delta\theta$ we could get with an index of refraction of 2.25 and a wavelength of 488 nm is 0.06° with $\theta_r = \theta_s = 45^\circ$.

Experimentally, we were unable to record good quality holograms with such oblique angles. Through trial-and-error, we settled on a geometry where the center ray of the signal and reference beams each made an angle of 35° with respect to the surface normal of the recording material. Theoretically, that gives a $\Delta\theta$ of 0.065° for the grating recorded between the center rays in the 1 mm thick LiNbO$_3$ crystal. This is roughly a 30% increase in the expected angular change required for Bragg-mismatching when compared to the results obtained for the 100 micron thick, photopolymer based holographic 3-D disk system.

To convert from $\Delta\theta$ to shift distance, we need to figure out where the focal point of the spherical reference beam is. For the photopolymer based holographic 3-D disk system, the focal point is placed just far enough from the backside of the disk so the reference illumination completely covers the signal beam spot on the photopolymer. For the 1 mm thick LiNbO$_3$ crystal, the entrance spot size is smaller than the exit spot size, since the signal beam expands inside the recording material. In order to cover the entire signal beam inside the recording material, the focal point of the spherical reference beam would have to be displaced further away from the recording material so the exit spot area is covered. However, that would significantly increase the amount of shift required to Bragg-mismatch a stored hologram. On the other hand, if the focal point of the spherical reference beam is placed so only the entrance spot is covered, then some of the signal illumination will not be recorded, leading to a weaker hologram. For this experiment we struck a compromise by covering all the entrance area and most of the exit area on the LiNbO$_3$. A decent hologram selectivity was maintained and the SNR of reconstructed holograms was sufficient for error free readout.

On a related topic, since the entrance spot size is different from the exit spot size, the intensity of the signal beam on the front face of the LiNbO$_3$ is different from the back surface. In order to get a good index modulation, the signal intensity should be similar to the reference intensity in the recording material. The intensities we finally settled on were 7.9 mW/cm^2 for the reference beam and 500 μW/cm^2 for the signal beam (both measured in the planewaves before the lenses). The intensity of the reference beam was a little weaker than the signal beam at the front surface and a little stronger than the signal beam at the back surface. This seems to produce the best quality holograms possible.

Other than just worrying about the possibility of packing enough holograms in the same volume to achieve a surface density of 100 bits/μm^2, we also have to make sure that the reconstructed holograms have enough power to be read out by the detector array. With the photopolymer based holographic 3-D disk, we did not have to worry about the diffraction efficiency of each hologram since the $M/\#$ is so high and only a few holograms overlapped in the same area. As a matter of fact, we consistently under-utilized the dynamic range of the photopolymer in order to simplify the exposure schedule and to record better quality holograms. With LiNbO$_3$, the $M/\#$ is usually much lower than the photopolymer given the same thickness. For the 1 mm thick iron doped LiNbO$_3$ crystal we used in this experiment, the $M/\#$ is 0.34 for vertically polarized incident beams and approximately 1 for horizontally polarized beams (this agrees with theory pretty well since the R$_{33}$ coefficient is about 3 times bigger than the R$_{13}$ coefficient in the electro-optic tensor for LiNbO3 [17]). So for horizontally polarized holograms, we can expect a diffraction efficiency that is 9 times higher than for holograms recorded with vertically

polarized beams (for the same exposure time). Therefore, we will use horizontally polarized reference and signal beams in this setup.

Even with a $M/\#$ of 1, the efficiency of the 1 mm thick $LiNbO_3$ still pales in comparison with DuPont's HRF-150-100 photopolymer. With $1/10^{th}$ the thickness, the $M/\#$ of the 100 micron thick photopolymer is around 6.5 times higher than the 1 mm thick $LiNbO_3$ (Chapter 2.4.4). In order to achieve a surface density of 100 bits/μm^2, we will need to store about 10 times more overlapping holograms in the same area than before. For a 1 mm thick photopolymer, this would not be a problem since the $M/\#$ would also increase by a factor of 10 due to the increased thickness. Therefore, the diffraction efficiency per hologram would remain about the same. However, for the 1 mm thick $LiNbO_3$, we will not have a $M/\#$ of 65 to play with. In terms of diffraction efficiency, each hologram recorded in the 1 mm thick $LiNbO_3$ will be 4,000 times less intense when compared to holograms recorded in a 1 mm thick photopolymer (if the full dynamic range of both materials were used). So for this experiment, we will really have to keep an eye on the SNR of the stored holograms to make sure the diffracted power is above the noise level. We will do this by monitoring the SNR as a function of surface density, from 1 hologram stored to 100 bits/μm^2. If at anytime the SNR drops below the desired value, we will modify the setup to either reduce the noise level or to increase the signal level.

To achieve the goal of 100 bits/μm^2 with the 1 mm thick $LiNbO_3$, we will first start with an in-track surface density of 50 bits/μm^2. This means the in-track holograms have to be separated by no more than 7.8 μm in horizontal shift. Then by overlapping one track at half the hologram height above and anther at half the hologram height below, boost the surface density to 100 bits/μm^2. It is kind of funny talking about the 'height' of

a hologram since the hologram has two different 'heights', one at the front face of the crystal and a bigger one at the back surface. We can go back and think about the surface density in terms of number of pixels in a hologram over the effective area of the hologram. For this experiment, we have again 590,000 pixels in each hologram over an effective hologram area of .75 mm (amount of vertical shift required to store another hologram) times 7.8 μm (amount of horizontal shift required to store the next in-track hologram), for a surface density of 100 bits/μm^2. Figure 3.55 shows the arrangement of holograms in the LiNbO$_3$ crystal.
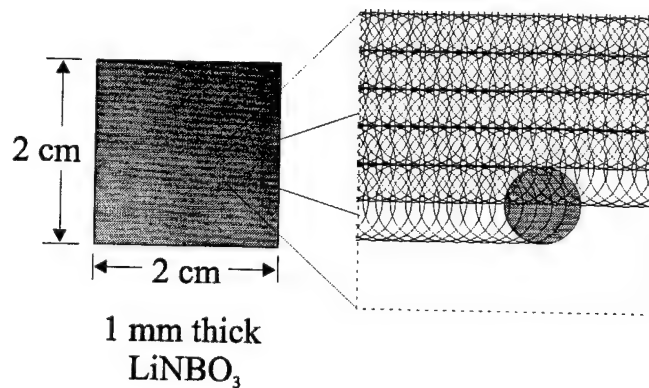

Figure 3.55 : The arrangement of holograms in the LiNbO$_3$ crystal.

3.4.3 Experimental Results

To determine if in-track holograms could be stored as closely as 7.8 μm in shift with the 1 mm thick LiNbO$_3$, we stored one hologram and measured its selectivity. Figure 3.56 shows the diffraction efficiency of the stored hologram as a function of shift distance. A 7.8 μm in-track shift between holograms corresponds to approximately the 2nd null for this setup. As explained previously in Chapter 3.4.2, the higher index of

refraction and the extra thickness of the LiNbO$_3$ hurt the selectivity a little. Instead of storing in-track holograms at the 4th null like in the holographic 3-D disk experiment (Chapter 3.3.5), for this experiment we will have to store additional in-track holograms at the 2nd null to get to 100 bits/μm^2. Therefore, more cross-talk noise is expected. For track overlapping, since the bandwidth of the signal and reference beams are roughly the same for the stored holograms, when the recording material is shifted radially by .75 mm, the reconstruction of the Bragg-matched strip is deflected completely off of the detector plane.
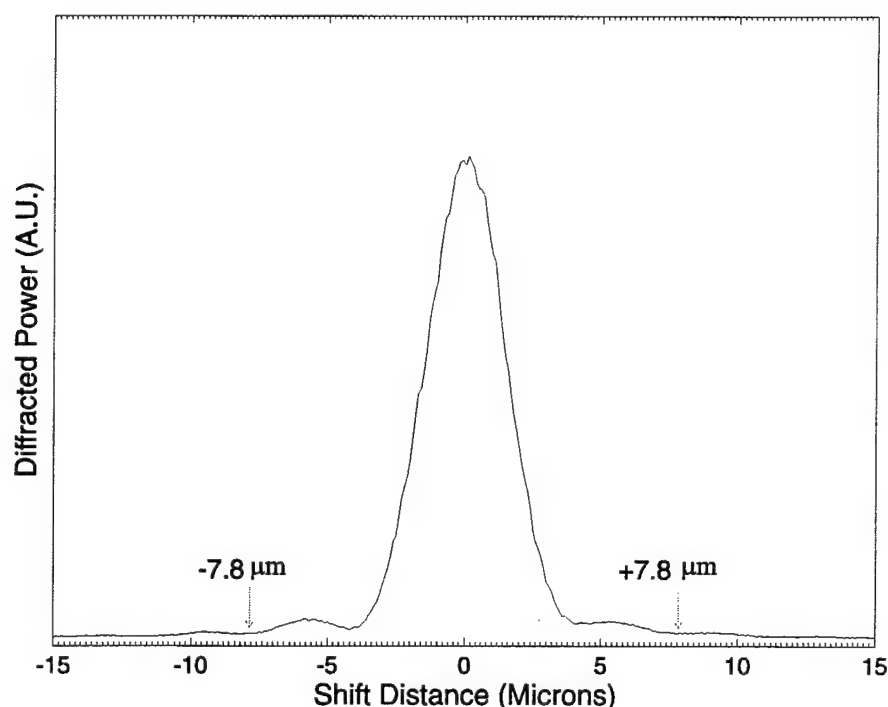

Figure 3.56 : Diffracted power as a function of shift distance for a high density hologram recorded with our shift multiplexing setup, using a 1 mm thick LiNbO$_3$.

We now know the setup can achieve a surface density of 100 bits/μm^2 as far as hologram selectivity is concerned. But before jumping directly to do the 100 bits/μm^2 experiment, we would like to take it step by step and solve possible difficulties along the way. Figure 3.57 shows the measured SNR of the system under various experimental

conditions. A SNR of 9.5 was measured by imaging the E-beam lithographed data mask

through the system without any recording material. Even though the imaging system for

the LiNbO$_3$ setup is exactly the same as the holographic 3-D disk experiment of Chapter

3.3.5, the system SNR is lower here. I suspect the mode quality of the Argon laser is not

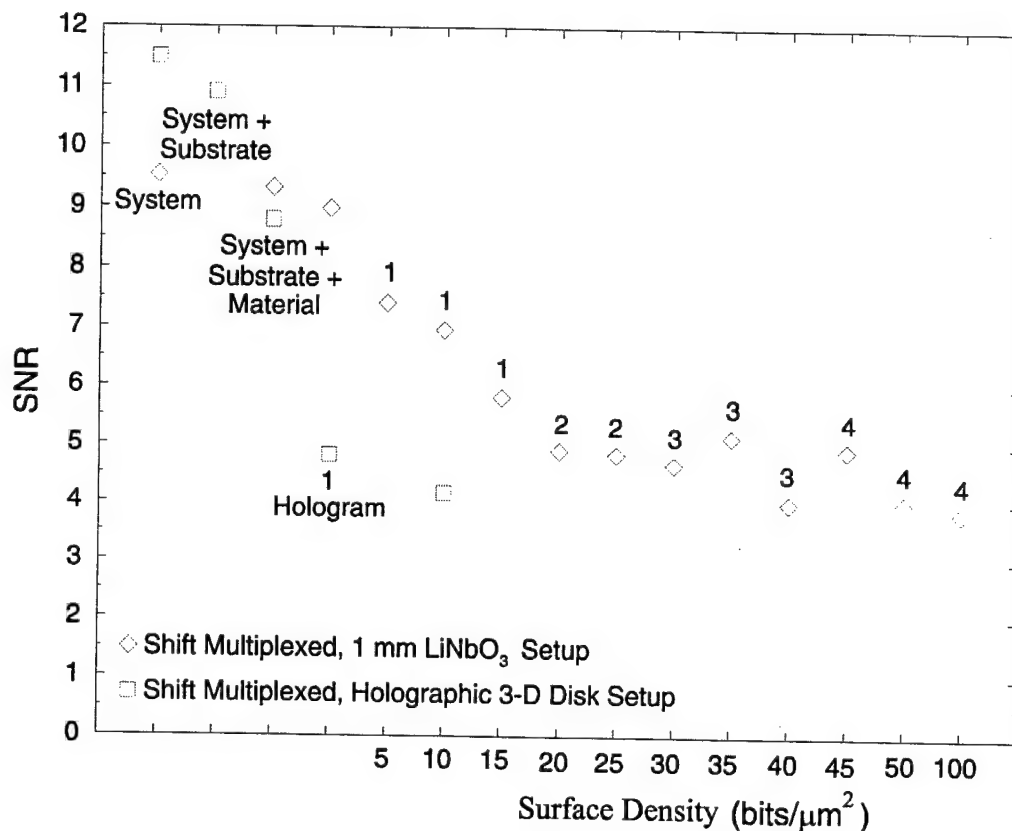quite as good as the DPSS laser used in the holographic 3-D disk experiment.

Figure 3.57 : SNR characterization for the shift multiplexed 1 mm LiNbO$_3$ setup.

When the AR coated 1 mm thick LiNbO$_3$ is placed at the recording plane, the

SNR drops slightly to 9.3. However, due to LiNbO$_3$'s much better optical quality and AR

coatings on both sides, the SNR of the LiNbO$_3$ setup becomes better than the

photopolymer based holographic 3-D disk setup. The difference in SNR between the two

setups is even larger for the case where the SNR of one hologram is analyzed. For the

LiNbO$_3$ setup, a SNR of 9 was obtained as compared to a SNR of 4.8 for the

photopolymer disk experiment. Again, the better optical quality and AR coatings really helped the LiNbO$_3$ setup in obtaining a higher SNR. To measure the effect of hologram multiplexing on the SNR, we stored shift multiplexed holograms in the LiNbO$_3$ at 5 bits/μm^2 in-track density increments, up to 50 bits/μm^2. For each experiment, the SNR of a reconstructed hologram near the middle of the track is measured. If the SNR is good and no errors are observed, then the in-track density for the next experiment is increased by 5 bits/μm^2 by decreasing the separation between adjacent holograms. However, if the SNR drops much below the SNR obtained for the 10 bits/μm^2 photopolymer based holographic 3-D disk experiment or if errors are observed, then parameters of the setup is modified to achieve a better SNR. For the final 100 bits/μm^2 experiment, the 50 bits/μm^2 in-track density experiment is repeated three times with a track separation of .75 mm. Table 3.1 shows a summary of the setup parameters and experimental results for the in-track density experiments with LiNbO$_3$.

For the 5, 10, and 15 bits/μm^2 in-track density experiments, the SNR measured were all well above 4.16 (the value measured for the 10 bits/μm^2 photopolymer based holographic 3-D disk experiment) and no errors were observed. Each hologram was recorded with a constant 20 seconds exposure and for the 5 bits/μm^2 experiment, the dynamic range of the LiNbO$_3$ was not yet fully utilized. Figure 3.58 shows the diffraction efficiency as a function of material shift for the 15 bits/μm^2 in-track density experiment. Notice the diffraction efficiency at the beginning and near the end is much higher than the middle (the power detector saturated for the last few holograms). This is caused by the erasure effect inherent in photorefractive materials such as LiNbO$_3$. Holograms multiplexed later in the sequence decay the strength of the already stored holograms.

Since the holograms at the beginning and the end of the track have fewer adjacent holograms, they don't get erased as much (for the last hologram, it doesn't get erased at all).

In-track Density (bits/μm^2)	# of Holograms	Hologram Separation (μm)	Exposure Time (seconds)	Average Diffraction Efficiency (%)	SNR
5	75	78	20	0.06	7.4
10	150	39	20	0.04	7
15	230	26	20	0.017	5.8
20	track 1 : 155 track 2 : 155	track 1 : 39 (0) track 1 : 39 (+19.5)	track 1 : 20 track 2 : 9	0.01	4.9
25	track 1 : 192 track 2 : 192	track 1 : 31.2 (0) track 2 : 31.2 (+15.6)	track 1 : 20 track 2 : 7	0.004	4.8
30	track 1 : 155 track 2 : 155 track 3 : 155	track 1 : 39 (0) track 2 : 39 (+13) track 3 : 39 (+26)	track 1 : 20 track 2 : 10 track 3 : 5	0.0038	4.7
35	track 1 : 175 track 2 : 175 track 3 : 175	track 1 : 34.2 (0) track 2 : 34.2 (+11.4) track 3 : 34.2 (+22.8)	track 1 : 20 track 2 : 8 track 3 : 5	0.0063	5.2
40	track 1 : 205 track 2 : 205 track 3 : 205	track 1 : 29.4 (0) track 2 : 29.4 (+9.8) track 3 : 29.4 (+19.6)	track 1 : 20 track 2 : 7 track 3 : 4	0.004	4
45	track 1 : 172 track 2 : 172 track 3 : 172 track 4 : 172	track 1 : 34.8 (0) track 2 : 34.8 (+8.7) track 3 : 34.8 (+17.4) track 4 : 34.8 (+26.1)	track 1 : 20 track 2 : 7 track 3 : 4 track 4 : 3	0.0052	4.9
50	track 1 : 192 track 2 : 192 track 3 : 192 track 4 : 192	track 1 : 31.2 (0) track 2 : 31.2 (+7.8) track 3 : 31.2 (+15.6) track 4 : 31.2 (+23.4)	track 1 : 20 track 2 : 6 track 3 : 4 track 4 : 3	0.0015	4

Table 3.1 : A summary of setup parameters and experimental results for the in-track density experiments with $LiNbO_3$.

The erasure effect will also force us to change the way in which the holograms are stored. For the 20 bits/μm^2 in-track density experiment, when all 310 holograms were stored sequentially by using the constant 20 seconds exposure per hologram, a lot of

errors were observed on the right side of the reconstructed holograms. The SNR for the left, center, and right sides of a reconstructed hologram were 5.5, 4.3, and 1.7, respectively. With a SNR of only 1.7 for the right side, the 'on' pixels were indistinguishable from the 'off' pixels. The reason why the SNR on the right side is so poor has to do with non-uniform erasure introduced by shift multiplexing in $LiNbO_3$.
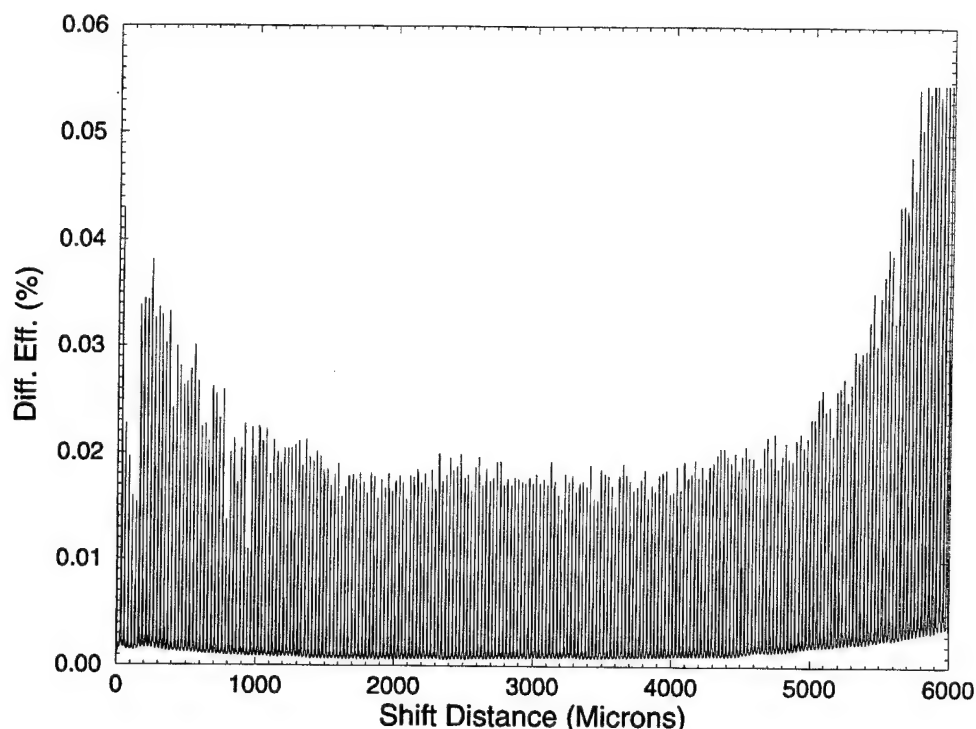

Figure 3.58 : Diffraction efficiency as a function of shift distance for the 15 bits/μm^2 in-track density experiment with $LiNbO_3$.

Figure 3.59 (a) shows a single track of holograms shift multiplexed sequentially from left to right and Figure 3.59 (b) shows the resulting diffraction efficiency. Once a hologram is stored, additional holograms are shift multiplexed to the right in partially overlapping areas. Therefore, the right side of a stored hologram becomes more erased, producing a reconstruction similar to the one shown in Figure 3.59 (a) (where the shaded black area represents lower intensity). This problem was also present in the 5, 10, and 15

bits/μm^2 in-track density experiments but it did not cause any errors until the 20 bits/μm^2 experiment (because more holograms are stored). We need to solve this non-uniform erasure problem before moving on to the next higher density experiment.
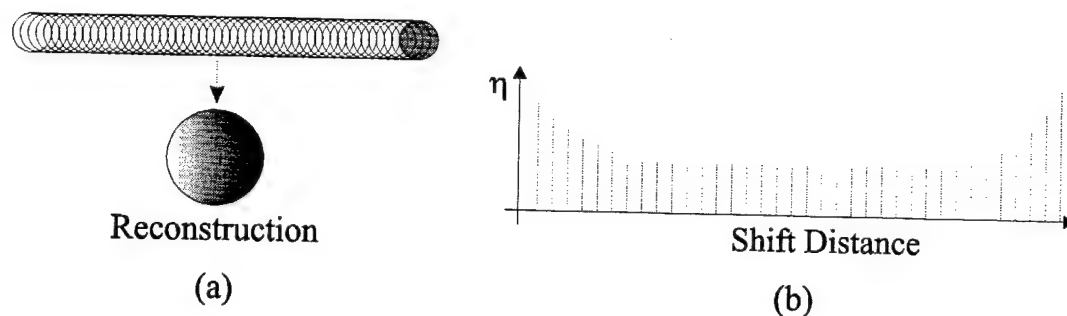

Reconstruction

(a)

η

Shift Distance

(b)

Figure 3.59 : Shift multiplexing in LiNbO$_3$ using a constant exposure schedule. (a) A single track shift multiplexed sequentially from the left to the right and a reconstruction from the middle of the track. (b) Diffraction efficiency of the stored holograms as a function of shift distance.

One method to solve the non-uniform erasure problem is to interleave the holograms during recording. Instead of storing all the holograms sequentially in a single pass, a track can be recorded by doing multiple passes as shown in Figure 3.60 (a). The starting point of each pass is shifted slightly so when all the holograms are recorded, the resulting track looks the same as Figure 3.59 (a). Since the holograms in each pass do not overlap, there is no non-uniform erasure. However, we do have to take into account the uniform erasure from multiple passes. The exposure schedule to achieve similar diffraction efficiency for multiple passes can be computed in the same matter as for angle multiplexed holograms [18]. Basically it involves recording the earlier passes stronger than the later passes, according to properties of the photorefractive recording material (Figure 3.60 (b)). Erasure from recording subsequent passes decays the strength of the earlier passes so they all have the same diffraction efficiency when the track is completed.

With this method of recording in-track holograms, the intensity and the SNR of the reconstruction should be more uniform across the entire hologram.
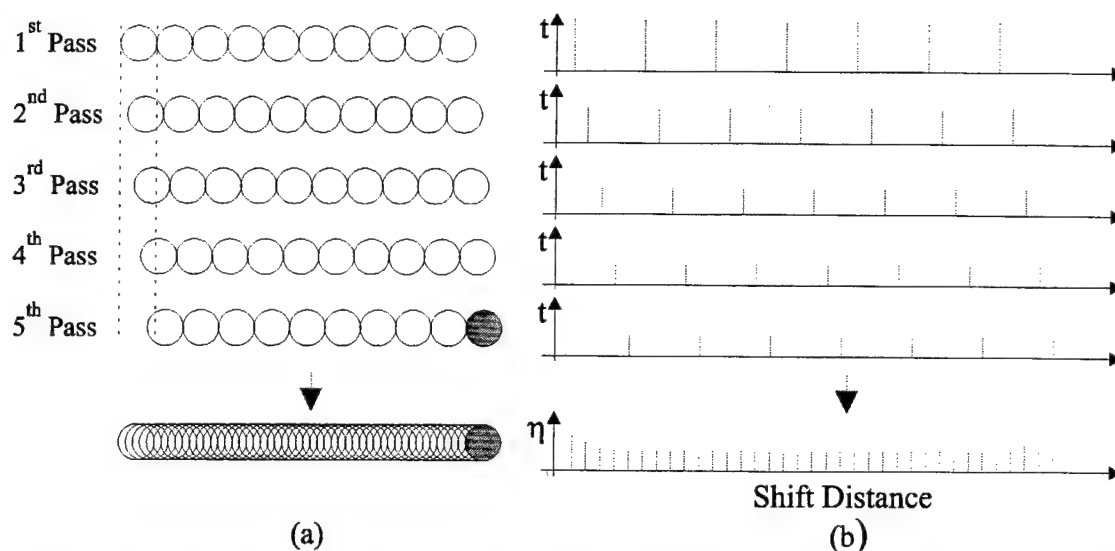

Figure 3.60 : Shift multiplexing in LiNbO$_3$ using multiple passes. (a) A single track shift multiplexed with multiple passes. (b) Exposure time for the different passes and the resulting diffraction efficiency of the stored track.

Unfortunately for our setup, it is not feasible to record completely non-overlapping holograms in each pass. The signal beam spot diameter on the entrance side of the LiNbO$_3$ is 1.5 mm. In order to separate the next hologram by 1.5 mm, we would have to wait approximately 5 minutes for the automated translation stage to move the recording material that far (we operated the KLINGER linear stage in the "slow" mode to obtain better repeatability). At that rate, the entire 20 bits/μm^2 in-track density experiment would take more than 24 hours. Instead, we compromised by recording multiple passes with partially overlapping holograms in each pass to save time. For example, the uniformity of the holograms in the 10 bits/μm^2 experiment was good enough to have zero errors. So for the 20 bits/μm^2 experiment, we recorded two 10 bits/μm^2 passes with the second pass displaced from the first by 19.5 μm. The holograms in the

first and second passes were recorded for 20 and 9 seconds each, respectively. These exposure times were obtained through trail-and-error to give the most uniform diffraction efficiency. The entire track took two hours to record and the resulting SNR was 4.9 with no errors observed.

For the 25 bits/μm^2 in-track density experiment, the track was recorded with two passes and the SNR was sufficient for error free reconstruction. However for the 30 bits/μm^2 experiment, the non-uniform erasure effect again caused errors and we had to go to three passes to record a single track without errors. To see if changing the modulation index would help improve the SNR, the reference and signal beam intensities were increased to 9.3 mW/cm^2 and 532 mW/cm^2, respectively (both measured from the planewaves before the lenses) for the next experiment. The SNR and the diffraction efficiency did improve slightly with the new modulation index for the 35 bits/μm^2 experiment. However for the 40 bits/μm^2 in-track density experiment, the SNR again dropped close to the error level due to non-uniform erasure. Figure 3.61 shows a portion of the diffraction efficiency as a function of material shift curve for the 40 bits/μm^2 experiment.

For the 45 bits/μm^2 experiment, the number of passes was increased to 4 and we began to notice significant cross-talk noise between adjacent holograms. A shift separation of 8.7 μm between holograms corresponds to a null spacing of approximately 2.2 according to Figure 3.56. The amount of cross-talk introduced was enough to cause some errors in the reconstruction. To decrease the inter-page cross-talk, we increased the selectivity of the setup by moving the focal point of the spherical reference beam closer to

the recording material. Figure 3.62 shows the resulting selectivity curve for a hologram recorded with the refocused spherical reference beam.
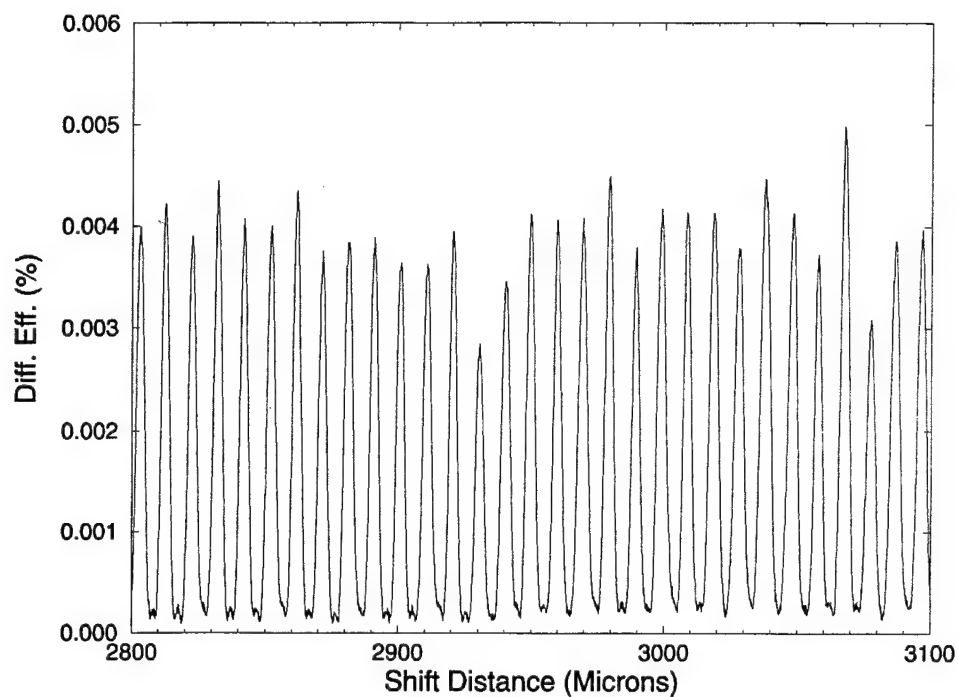

Figure 3.61 : A portion of the diffraction efficiency as a function of shift distance curve for the 40 bits/μm^2 in track density experiment with LiNbO$_3$.
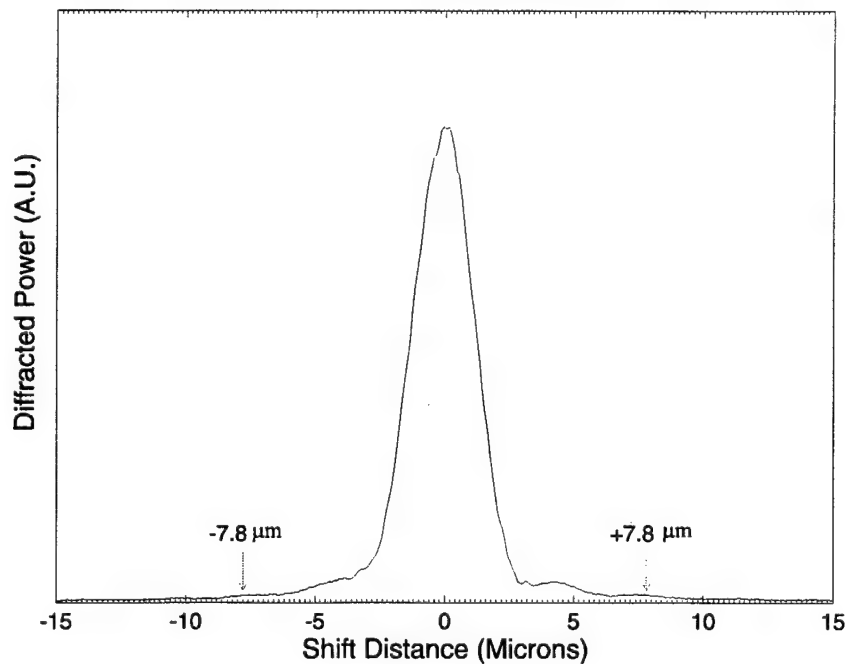
Figure 3.62 : Diffracted power as a function of shift distance for a high density hologram recorded with the refocused spherical reference beam.

With the refocused reference beam, a hologram separation of 8.7 μm now corresponds to a null spacing of 2.8 for the 45 bits/μm^2 experiment. Furthermore, for the 50 bits/μm^2 in-track density experiment, adjacent holograms can be stored at 2.5[th] null. The better selectivity should decrease inter-page cross-talk noise and increase the SNR of the stored holograms. However, by moving the focal point of the spherical reference beam closer to the recording material, less of the signal spot at the exit side of the $LiNbO_3$ is covered. This might decrease the hologram SNR. For the 45 bits/μm^2 experiment with 4 passes and refocused reference beam, the SNR obtained was 4.9 and no errors were observed. This shows that the gain in inter-page cross-talk noise suppression out-weighted the loss of SNR due to a smaller recording volume. The average diffraction efficiency also showed a slight improvement, possibly due to the increased reference beam intensity during recording.

The 50 bits/μm^2 in-track density experiment presented us with a new challenge. For the first trial with 4 passes, the resulting SNR for the left, center, and right sides were 3.5, 3.6, and 2.4, respectively. The average diffraction efficiency was 2.5×10^{-5} and there were some errors in the reconstruction. The chief cause of these errors was the scattering noise overpowering the reconstructed hologram. The diffraction efficiency of this experiment was dangerously close to the scattering diffraction efficiency of $\eta_{scat} \simeq 10^{-5}$. η_{scat} was determined by passing the reference beam through the 1 mm thick $LiNbO_3$, measure the total energy collected by the Nikon lens at the detector plane, and then normalize it by the reference beam power. For this setup, nearly all the scattering noise

was generated by the reference beam lens since η_{scat} did not change appreciably with the LiNbO$_3$ removed. Reflections between the multiple optical elements in the Universal $f/1.1$ projection lens create a cone of scattered light with the intensity strongest in the direction of the reference beam. The Nikon lens captures some of the scattering and images it to the detector plane. Since the right side of the Nikon lens is closer to the reference beam, the right side of the detector plane has more scattering noise. This combined with weaker diffraction efficiency on the right sides of the stored holograms caused most of the errors in the 50 bits/μm^2 experiment.

There are two immediate solutions to the scattering problem : (1) increase the diffraction efficiency of the stored holograms; (2) reduce the scattering diffraction efficiency. Since we are already using the entire dynamic range of the LiNbO$_3$, it is not possible to raise the diffraction efficiency much higher than the current level. We could make the diffraction efficiency more uniform across the hologram by making more passes per track, but that would only be a temporary solution since the diffraction efficiency of each hologram will still be pretty low. Therefore, we have to reduce the scattering noise level through some type of filtering process in order to increase the SNR. Figure 3.63 shows the setup we used to filter the scattering noise for the 50 bits/μm^2 in-track density experiment. The new setup is exactly the same as the one shown in Figure 3.52, except for a second 4-f system in the signal arm. A spatial filter with an opening just big enough to pass the signal beam is placed at the Fourier plane of the second 4-f system. Since scattering originates from a different spatial position than the signal beam, most of it becomes blocked by the spatial filter. This is evident in the much lower scattering diffraction efficiency of $\eta_{scat} \simeq 3.4 \times 10^{-7}$ for the new setup. However, with more lenses

in the signal beam, the system SNR might drop. We couldn't measure the system SNR directly since we only had three Nikon *f*/1.4 camera lenses, but we could measure the SNR of a single hologram recorded with the new setup and it was 7.6. Compared with a SNR of 9 for a single hologram recorded with the setup shown in Figure 3.52, we lost 1.4 in SNR due to the extra pair of Nikon lenses in the signal beam. So the question now is: for low diffraction efficiency holograms, do we gain more SNR from the filtering process than we lose on the extra 4-*f* system? To find out, the same 50 bits/μm^2 in-track density experiment was repeated with the new setup and a SNR of 4 was obtained. This SNR is better than without filtering and no errors were observed in the sampled reconstruction. The measured diffraction efficiency of the stored holograms did go down to 1.5×10^{-5} since most of the scattering noise was blocked.
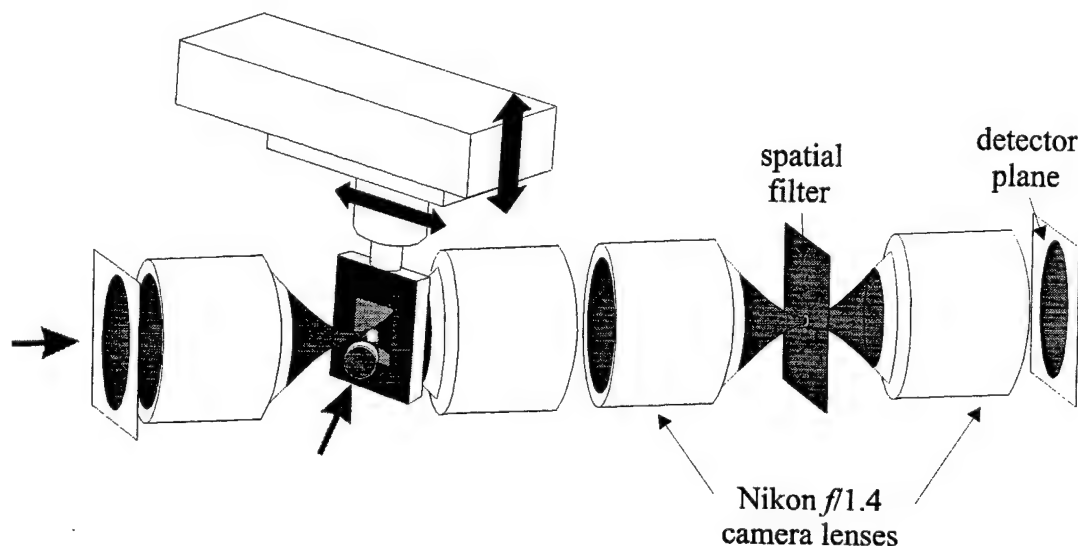

Figure 3.63 : Setup for the 50 and 100 bits/μm^2 high density experiments (with filtering).

A surface density of 100 bits/μm^2 was achieved by repeating the 50 bits/μm^2 in-track density experiment three times in partially overlapping tracks. Each track was separated by a vertical displacement of .75 mm and the in-track holograms were shifted

horizontally by 7.8 µm. This gave an effective area per hologram of .75 mm × 7.8 µm = 5850 µm². Since each hologram contained 590,000 pixels, the surface density achieved was 590,000 bits / 5850 µm² = 100 bits/µm². Figure 3.64 shows a portion of the diffraction efficiency as a function of shift distance curve for the middle track of the 100 bits/µm² experiment. The holograms were stored very close together so inter-page cross-talk was the source of most of the noise encountered. The tracks were recorded sequentially from top to bottom and the average diffraction efficiency of the holograms in the middle track was 0.00044%. This diffraction efficiency is much lower than the previous 50 bit/µm² in-track density experiment because the bottom track partially erases the middle track. Nevertheless it is about ten times higher than the scattering diffraction efficiency after filtering. The SNR achieved was 3.85 and no errors were observed in the sampled reconstruction when localized threshold values were used (since the intensity of the reconstructed hologram varied across the page due to the non-uniform erasure effect).
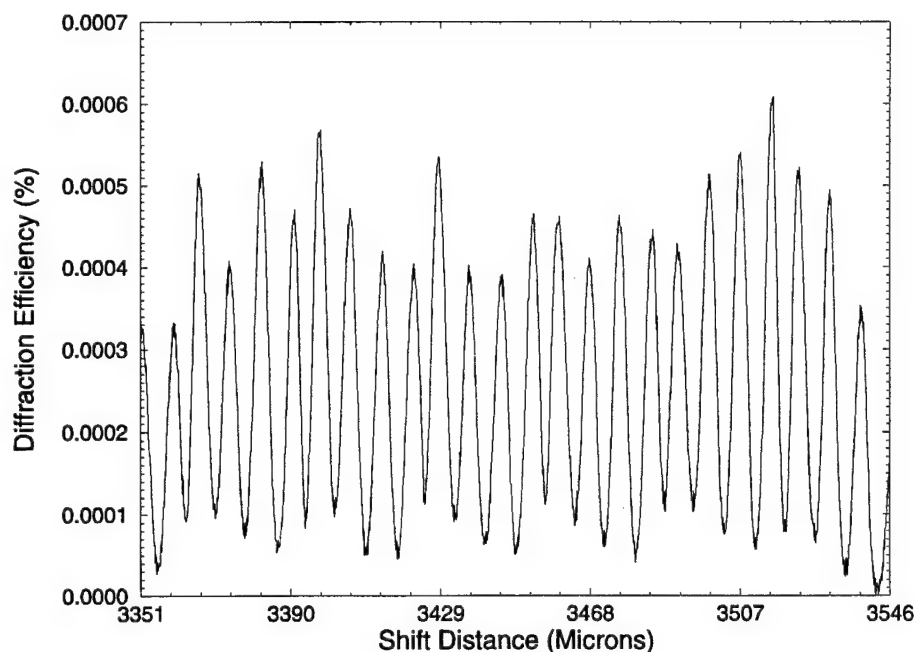
Figure 3.64 : A portion of the diffraction efficiency as a function of shift distance curve for the middle track of the 100 bits/μm^2 experiment with LiNbO$_3$.

Figure 3.65 shows a reconstructed hologram from the middle of the middle track. The three frames correspond to the left, center and right sides of the reconstructed hologram. The visual quality of these frames are slightly poorer than the images shown in Figure 3.48 for the 10 bits/μm^2 photopolymer based 3-D disk experiment. To estimate the bit-error-rate (BER) of 100 bits/μm^2 experiment, a histogram was generated from the reconstructed hologram and then fitted to a first order χ^2 distribution (as shown in Figure 3.66). The estimated BER obtained by integrating the tail areas of the χ^2 distribution for both the 'on' and 'off' pixels was approximately 9.5×10^{-5}. This BER is about the same as the original 10 bits/μm^2 surface density experiment using peristrophic and angle multiplexing.
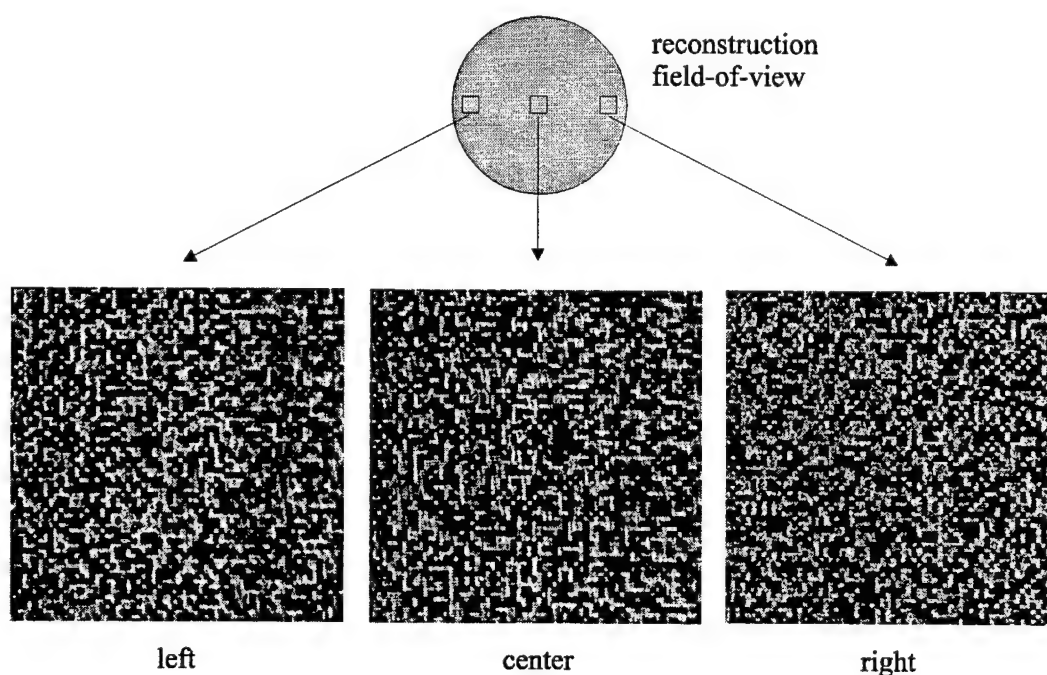

Figure 3.65 : The left, center, and right sides of a reconstructed hologram for the 100 bits/μm^2 experiment using LiNbO$_3$.

In Figure 3.66, the histograms for the 'on' and 'off' pixels actually cross-over. This means that a global intensity threshold value cannot be picked to classify all the pixels correctly for the entire reconstruction. Due to the non-uniform erasure effect from shift multiplexing partially overlapping holograms in LiNbO$_3$, we need to use localized threshold values to classify the recalled pixels. When different threshold values were used to classify the left, center, and right sides, the reconstructed pixels from a sampled hologram were all classified correctly.
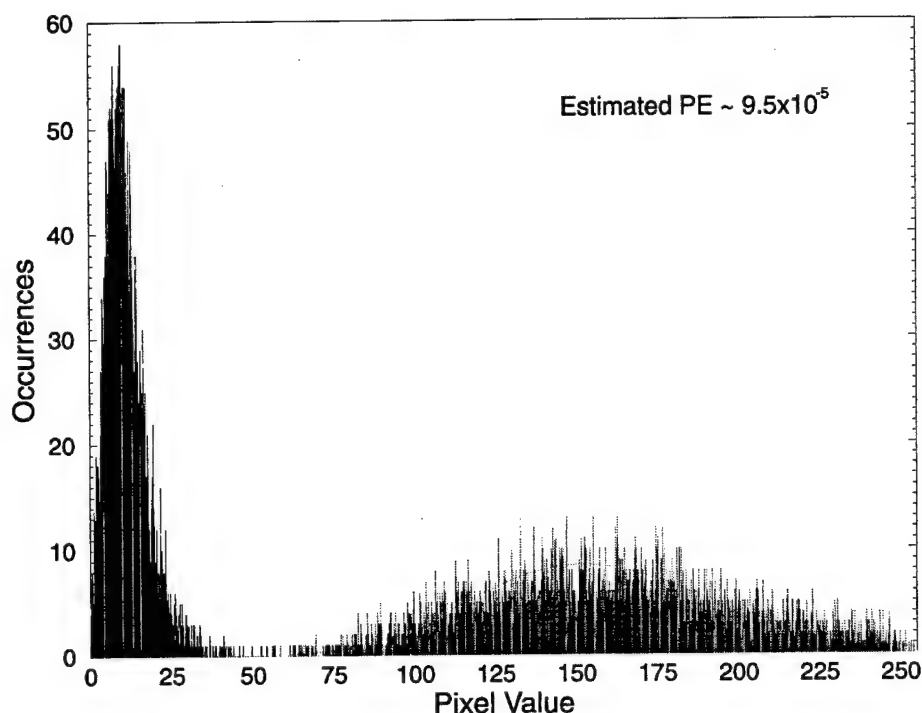
Figure 3.66 : Combined histogram obtained from three different sampled windows for the 100 bits/μm^2 experiment using LiNbO$_3$.

3.5 Discussions and Conclusions

The market segment where we think a holographic data storage device can achieve acceptance, with currently available components, is mass distribution of pre-

mastered information. This brings the holographic data storage device into direct competition with the widely popular compact disc. In order to succeed in a market where the compact disc holds dominance, the holographic data storage device must offer better performance at a comparable price. In this chapter we have demonstrated that a surface density many times higher than the compact disc can be achieved holographically. One of the systems implemented used a photopolymer based holographic 3-D disk that has the same physical dimensions as the compact disc. The holograms were recorded sequentially in partially overlapping areas along a spiral track. A spherical reference beam was used to reconstruct the stored holograms while the rotational motion of the disk provided access to different holograms. This system was designed to be similar to the compact disc so many existing components could be used to reduce cost. Table 3.2 summarizes the experimental results we have obtained so far and the target values we need to make holographic 3-D disks competitive.

	100 μm Thick Photopolymer Based Holographic 3-D Disk System	1 mm Thick $LiNbO_3$ System	Target System
Surface Density	10 bits/μm^2	100 bits/μm^2	100 bits/μm^2
Capacity per 120 mm Disk	10 Gbytes	100 Gbytes	100 Gbytes
Recording Rate	1.1 Mbits/sec	70 Kbits/sec	as fast as possible
Readout Transfer Rate	.4 Mbits/sec	.4 Mbits/sec	> 100 Mbits/sec
Average Diffraction Efficiency	0.5%	0.00044%	as high as possible
Raw Bit-Error-Rate	2.4×10^{-5}	9.5×10^{-5}	$<10^{-6}$

Table 3.2 : Experimental results obtained so far and the target values.

For the 100 micron thick photopolymer based holographic 3-D disk system, we achieved a surface density of 10 bits/μm^2. The total data capacity per 120 mm disk at 10 bits/μm^2 is about 10 Gbytes, roughly ten times higher than the conventional compact disc. However, the projected surface density of the next generation compact disc, the Digital-Versatile-Disk (DVD), is around 6 bits/μm^2 per layer. The single layered DVD will be introduced in the first quarter of 1997 and the dual layered disks sometime shortly after. The maximum number of layers planned so far is 4, which brings the total surface density of the most advanced DVD up to approximately 20 bits/μm^2. In order for holographic 3-D disks to be competitive, we need to have a surface density that is much higher than the maximum projected density of DVD. A simple way to increase the surface density of holographic 3-D disks is to use a thicker recording material. To reach the target surface density of 100 bits/μm^2, we would need a 1 mm thick photopolymer. However, a 1 mm thick photopolymer is not currently available so we used a 1 mm thick LiNbO$_3$ to achieve 100 bits/μm^2. For a 1 mm thick holographic 3-D disk with a surface density of 100 bits/μm^2, the capacity per 120 mm disk is around 100 Gbytes, roughly 5 times higher than the projected capacity of the most advanced DVD.

One major performance issue we have neglected completely in the high density experiments is data transfer rate. For a read only holographic 3-D disk system, the recording rate is probably not a big issue. Nevertheless the recording material should be sensitive enough so the disks can be replicated quickly. The readout transfer rate on the other hand is a key measure of system performance. Holographic data storage devices have the potential for very fast data readout rate due to the page format of the stored holograms. For both the photopolymer disk and LiNbO$_3$ experiments, the CCD array

used was much smaller than the aperture of the Nikon lens. Approximately 140 × 100 data mask pixels were captured within the CCD array aperture at a frame transfer rate of 30 Hz. This gives a data readout rate of only .4 Mbits/s as compared to the projected DVD channel bit rate of about 25 Mbits/s. If we had a CCD array large enough to capture all 590,000 reconstructed pixels within 33 ms (30 Hz), then the data readout rate would be 17 Mbits/s, which is still slower than DVD. In order to achieve the target value of >100 Mbits/s data readout rate, we need an imager that can transfer 590,000 pixels at a frame rate of ~180 frames/s. Not an impossible task considering Kodak's EktaPro Model 4540 CCD [19] camera can achieve a maximum transfer rate of 4,500 frames a second. At 256 × 256 pixels per frame, that is about 280 Mbits/s data transfer rate. However, at such a high frame rate, the diffraction efficiency of the stored holograms becomes an issue. For a signal-to-detector noise-ratio of greater than 40:1 with most CCD detector arrays, it is necessary to collect approximately 1,000 photons per pixel. The amount of time it takes to accumulate P photons per pixel can be expressed as:

$$\tau = \frac{PhcN_p}{\eta I_{inc} \lambda} \tag{3.22}$$

where h is the Planck's constant, c is the speed of light, N_p is the total number of pixels, η is the diffraction efficiency, I_{inc} is the incident power of the reference beam, and λ is the wavelength. For example, if we chose the values from the 10 bits/μm^2 photopolymer disk experiment ($N_p = 590,000$, $\eta = 0.5\%$, $\lambda = 532$ nm, and a reference power of $I_{inc} = 20$ mW to read the stored holograms), then the required integration time per pixel, τ, is equal to ~ 2 μs, much less than the frame integration time of 200 μs at 4,500 frames/s. So with the

photopolymer as the recording material, we probably don't have to worry too much about collecting enough photons per pixel to achieve good signal-to-detector noise-ratio.

Another important measure of system performance is the raw bit-error-rate. For the high density experiments shown in this chapter, the estimated raw BER were in the range of 10^{-4} to 10^{-5}, similar to the raw BER of compact discs. Using error correction codes with similar overheads as the compact discs, the BER of the holographic systems can be made to be less than 10^{-12}. We can achieve the target raw BER of 10^{-6} by eliminating more noise sources such as multiple reflections and reference beam scattering through AR coating.

What we ultimately want to construct is a holographic 3-D disk storage device similar to the one shown in Figure 3.67. The system consists of two major components : the read/write head and the holographic 3-D disk. The read/write head is mounted on a mechanical slide so the spherical reference beam can access different tracks on the disk while the spinning motion of the disk continuously reconstructs the in-track holograms. The spherical reference beam reconstructs a page of data at a time and the data is readout with a 2-dimensional detector array. For a system with recording capability, the data is presented on the SLM and then recorded on the disk by mixing it with the spherical reference beam. For a read-only drive, the signal beam is not required and the system becomes very similar to compact disk drives. Notice there are no moving parts inside the read/write head and only about ten components are required. Some standard components from compact disc drives, such as the objective lens assembly, can be used in this setup to generate the spherical reference beam and to provide focus and tracking control. The

holographic 3-D disk setup shown in Figure 3.67 offers both the simple compact disc-like architecture and high speed page access/high surface density of a holographic device.
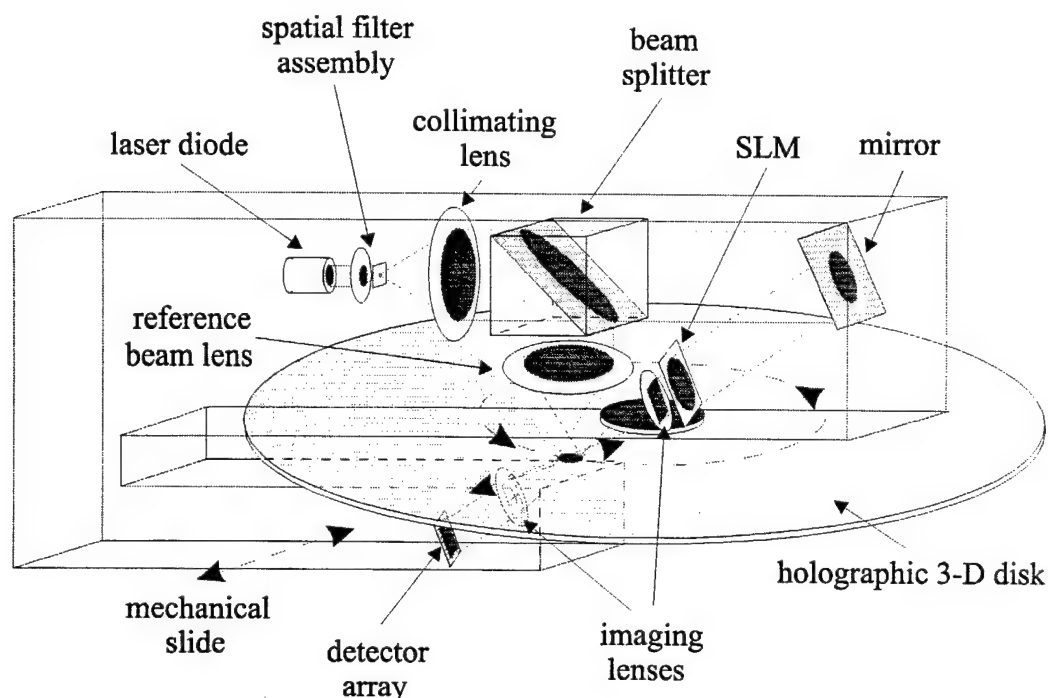

Figure 3.67 : A possible compact implementation of a photopolymer based holographic 3-D disk system.

We are still a fair distance away from making a holographic 3-D disk system that can be marketed as a commercial product. The biggest roadblock is the recording material. Currently, the most promising material for the type of application we have in mind is DuPont's photopolymer. However, in order to achieve the target surface density of 100 bits/μm^2, a recording material around 1 mm thick is required. So far, no 1 mm thick photopolymer is commercially available. Furthermore, similar to a compact disc drive, a fully operational holographic 3-D disk device has to provide focusing and tracking error signals. For the system shown in Figure 3.67, the spherical reference beam is on-axis with respect to the disk. Therefore, the same differential detection method used in compact disc drives can be applied here. However, if we want to use the shrinkage

insensitive recording geometry with the reference beam off-axis, then perhaps special markers could be embedded in the holograms to provide the error signals. Maybe a completely new method can be invented for holographic 3-D disk devices that exploits some aspect of holography. Further research and development is required to deal with the disk's axial and radial run-out problems.

To achieve the target readout transfer rate of greater than 100 Mbits/s, we could use fast CCD detectors such as Kodak's EktaPro cameras. However, the cost of these cameras exceeds $100,000 per unit and therefore would not be practical for our application. What is needed are CMOS detectors custom designed for holographic data storage. The CMOS detectors can be designed with multiple outputs for fast frame transfer and smart electronics to classify the reconstructed pixels on the fly. Mass production of these detectors should bring the cost down to be compatible with consumer electronics.

For the compact disc, a lot of research has been done in the area of data structure, modulation, and error-correction-codes to optimize the performance. With holographic 3-D disks, the data is stored as discrete 2-dimensional pages. The extra dimension might offer designers more flexibility (or headaches) in the selection of optimal codes. Some researchers are beginning to address these issues for holographic data storage [20-22] and a standard might emerge in the near future.

One of the reasons why the compact disc has been so successful is the ease of disk replication. A compact disc can be made in less than a minute for less than 50 cents. This has allowed the music industry to rake in huge profits in recent years. Fortunately, holographic 3-D disks can also be replicated quickly and cheaply. One possible process

involves copying the entire master disk through a single exposures [23]. This technique has been demonstrated for conventionally multiplexed holograms (angle, peristrophic, etc.) but further development is needed for shift multiplexed holograms.

In this chapter, a new multiplexing method that is especially well suited for the holographic 3-D disk application is shown. In its current implementation, a spherical lens in placed in the reference beam to generate a spherical reference beam for shift multiplexing. This technique works fine but perhaps the optimal shift multiplexing configuration is still undiscovered. It may be possible to obtain better shift selectivity and/or less inter-page cross-talk by making simple modifications to the reference beam. For example, Figure 3.68 shows the shift multiplexing selectivity curves for three different configurations. The solid black line is the same curve as Figure 3.38 while the two dotted curves were obtained by placing random phase plates of different resolutions in front of the spherical lens in the reference arm. By decreasing the pixel size of the random phase plate, the rays in the reference beam become more distorted and therefore the selectivity improves (at the expense of increased background noise level). It might also be possible to shape the intensity profile of the reference beam to reduce inter-page cross-talk noise [24].

It is hard to predict the future of holographic data storage base on current stage of development. Some have estimated its chances at 50/50 and I believe that is overly optimistic. We have presented most of the technical roadblocks but perhaps the biggest one of all is politics. Even if holographic 3-D disks can produce the target performance numbers, why should the compact disc industry abandon their proven technology?

Luckily, within a few years we will be able to tell if holographic data storage is a boom or bust.
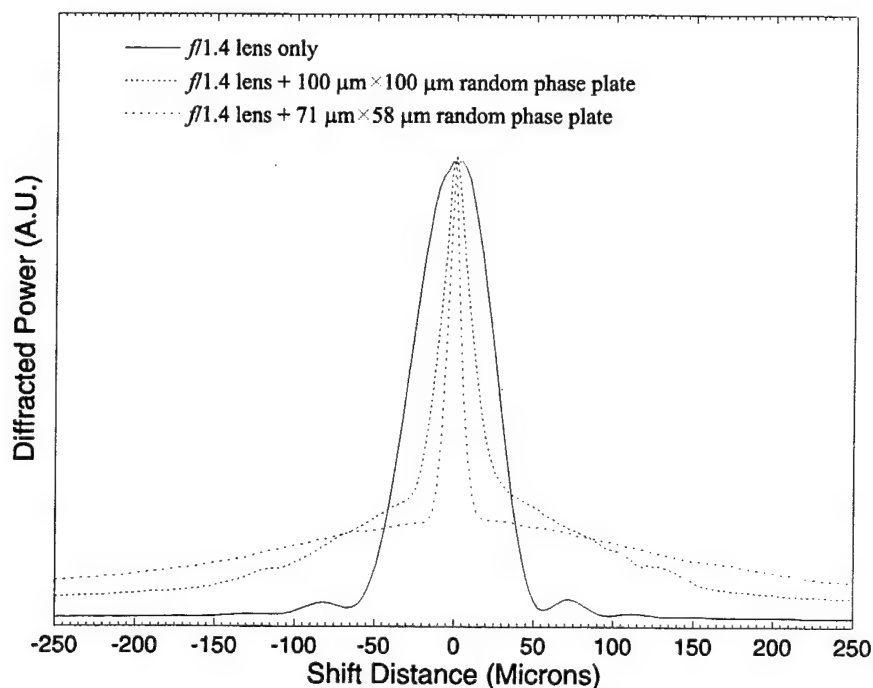

Figure 3.68 : Shift selectivity curves for three different configurations.

References

[1] D. L. Staebler, J. J. Amodei, and W. Philips. "Multiple storage of thick holograms in LiNbO₃," *VII International Quantum Electronics Conference*, Montreal, Canada (1972).

[2] F. H. Mok, M. C. Tackitt, and H. M. Stoll. "Storage of 500 high-resolution holograms in a LiNbO₃ crystal," *Optics Letters*, 16(8):605-607, 1991.

[3] S. Yin, H. Zhou, F. Zhao, M. Wen, Y. Zang, J. Zang, and F. T. S. Yu. "Wavelength-multiplexed holographic storage in a sensitive photorefractive crystal using a visible-light," *Optics Communications*, 101(5-6):171-176, 1991.

[4] G. A. Rakuljic, V. Levya, and A. Yariv. "Optical data storage by using orthogonal wavelength-multiplexed volume holograms," *Optics Letters*, 17(2):1471-1473, 1992.

[5] C. Denz, G. Pauliat, and F. Rosen. "Volume hologram multiplexing using a deterministic phase encoding method," *Optics Communications*, 85:171-176, 1991.

[6] Y. Taketomi, J. E. Ford, H. Sasaki, J. Ma, Y. Fainman, and S. H. Lee. "Incremental recording for photorefractive hologram multiplexing," *Optics Letters,* 16(22):1774-1776, 1991.

[7] D. Psaltis, D. Brady, X. G. Gu, and S. Lin. "Holography in artificial neural networks," *Nature*, 343(6256):325-330, 1990.

[8] F. H. Mok. "Angle-multiplexed storage of 5000 holograms in lithium niobate," *Optics Letters*, 18(11):915-917, 1993.

[9] K. Curtis, A. Pu, and D. Psaltis. "Method for holographic storage using peristrophic multiplexing," *Optics Letters*, 19(13):993-994, 1994.

[10] D. Psaltis, M. Levene, A. Pu, G. Barbastathis, and K. Curtis, "Holographic storage using shift multiplexing," *Optics Letters*, 20(7):782-784, 1995.

[11] *Toshiba SD (Super Density),* Format technical specifications. (Toshiba, Japan, 1995).

[12] H. Y. S. Li and D. Psaltis. "Three-dimensional holographic disk," *Applied Optics*, 33(17):3764-3774, 1994.

[13] C. B. Burckhardt. "Use of a random phase mask for the recording of Fourier transform holograms of data masks," *Applied Optics*, 9(3):695-700, 1970.

[14] Y. Takeda, Y. Oshida, and Y. Miyamura. "Random phase shifter for Fourier transformed holograms," *Applied Optics*, 11(4):818-822, 1972.

[15] A. Iwamoto. "Artificial diffuser for Fourier transform hologram recording," *Applied Optics*, 19(2):215-221, 1980.

[16] H. -Y. S. Li. *Photorefractive 3-D disks for optical data storage and artificial neural networks*. Ph.D. thesis, California Institute of Technology, 1994.

[17] A. Yariv. *Optical Electronics*, (Saunders College Publishing, Florida, 1991), Fourth Edition.

[18] D. Psaltis, D. Brady, and K. Wagner. "Adaptive optical networks using photorefractive crystals," *Applied Optics*, 27(9):1752-1759, 1988.

[19] *Kodak EktaPro HS Motion Analyzer, Model 4540*, Technical specifications. (Kodak, San Diego, 1995).

[20] J. F. Heanue, K. Gurkan, L. Hesselink. "Signal-detection for page access optical memories with intersymbol interference," *Applied Optics*, 35(14):2431-2438, 1996.

[21] W. F. Hau and A. A. Sawchuk. "Performance analysis of smart pixel interfaces for optical page-oriented memories," OSA annual meeting, MKK3, Rochester, New York. Oct 1996.

[22] M. A. Neifeld and J. D. Hayes. "Error-correction schemes for volume optical memories," *Applied Optics*, 34(35):8183-8191, 1995.

[23] F. Mok, G. Zhou, A. Chugh. "Read-only data storage system," OSA annual meeting, MAAA8, Rochester, New York. Oct 1996.

[24] M. A. Neifeld and M. McDonald. "Technique for controlling cross-talk noise in volume holography," *Optics Letters*, 21(16):1298-1300, 1996.

Chapter 4 : Optical Correlators Using Holographic Database

4.1 Introduction

Optical correlators have been around since as early as 1961 in the form of *hard-clipped filters* [1]. These frequency-plane filters were computed digitally and then placed in an optical system for radar data processing. Two years later, A. B. Vander Lugt demonstrated a new technique for synthesizing the frequency-plane filters optically [2], eliminating the complex task of computing the filters. The desired frequency-plane filter was synthesized with a single lens and then recorded onto a photographic plate with an interferometric system. To compute the correlation function, a lens calculates and projects the Fourier transform of the input image onto the frequency-plane filter. The frequency-plane filter modifies the phase and amplitude of the incident light to produce the cross-correlation at the output plane.

An array of these frequency-plane filters can be multiplexed at a single location on the photographic plate by varying the wavelength of the illumination source or by changing the carrier frequency of the interferometric system (like angle multiplexing). Since the filters are all stored at the same location on the photographic plate, an input image can be cross-correlated with all the filters simultaneously to produce an array of

correlation peaks. The number of different filters that can be stored at the same location is limited by the dynamic range of the recording material and the number of wavelengths or carrier frequencies that can be introduced in the system. Vander Lugt was able to store up to nine different filters at a single location using photographic plates [3].

The advantage of an optical correlator system as compared to its electronic counterpart has always been speed. An optical correlator system can essentially compute 2-dimensional cross-correlations between an input image and the frequency-plane filters at the speed of light. In most optical correlator systems, the computation rate is limited only by the spatial-light-modulator (SLM) update rate and detectors, which can range from tens of Hertz to thousands of Hertz. Furthermore, Vander Lugt correlators require no external memory or supervision to supply the frequency-plane filters. The filters are already embedded in the system as one of the optical elements (i.e. filters stored in photographic plates). Vander Lugt correlators can be thought of as a self-contained black-box. It computes the cross-correlation between an input image and all the stored filters simultaneously and produces an array of correlation peaks at the output plane.

Even though optical correlators have been around since the early 60's, they have not found a significant number of applications in non-defense related fields. The problems with optical correlators had mainly been on how to introduce the input images into the optical correlator and how to read the results of the cross-correlation. These operations used to require special electro-optical elements that were very expensive. That has changed considerably in the past decade. The consumer electronics boom of the 80's introduced a wide range of affordable components that can be used to interface with optical correlators. For example, liquid crystal display (LCD) panels and micro-mirror

arrays found in projection televisions can be used as input devices in optical correlators. The resolution of currently available LCD panels range from 256×256 to 1024×1024 and some can be updated at frame rates greater than 30 Hz. For reading the correlation output, cheap CCD arrays from video camcorder can be used. The intensity and position of the correlation peaks can be determined with a personal computer analyzing the signal from the CCD array.

Other important advances such as laser diodes, high dynamic range recording materials (such as DuPont's photopolymer), new holographic multiplexing techniques and better understanding of holographic data storage have allowed us to construct small and relatively inexpensive Vander Lugt type optical correlator systems with a large database of filters. By programming the database with appropriate filters, we have used these systems to navigate an autonomous vehicle, to recognize/track targets, and to identify users through sequential fingerprint inputs, all in real-time. In this chapter, results from those experiments will be shown.

4.1.1 Optical Correlator Basics

We are interested in Vander Lugt optical correlator systems because they are very similar to the normal Fourier plane holographic data storage setup shown in Figure 4.1 (a). The data to be stored is presented on the SLM and imaged to the detector array with a 4-f lens system. The recording material is placed at the exact Fourier plane of the 4-f system, normal to the signal beam. A hologram is recorded by mixing the Fourier transform of the input data pattern with a planewave reference beam. To recall the stored

data, the reference beam is re-introduced and the signal beam is reconstructed. A detector

array then reads the reconstructed hologram and reproduces the original data.
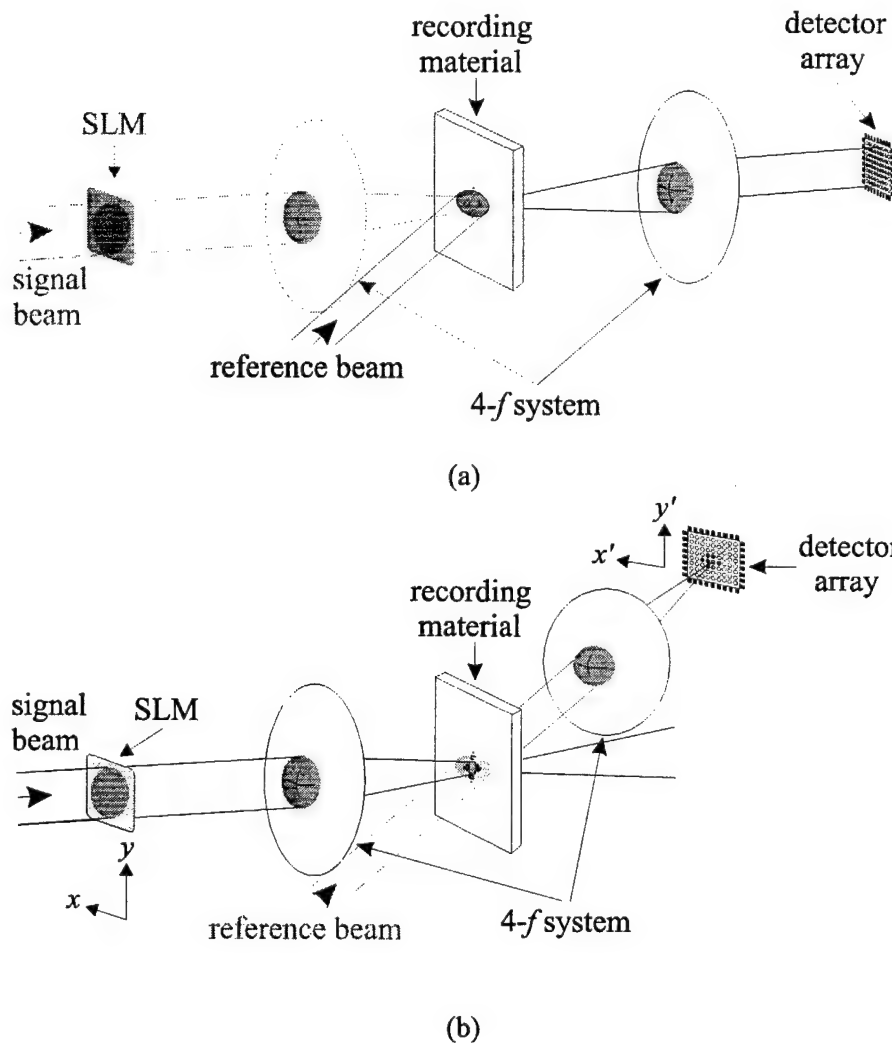

(a)

(b)

Figure 4.1 : Holographic optical memory and correlator. (a) Fourier plane holographic memory setup. (b) Vander Lugt type optical correlator setup using holographic filters.

For a Vander Lugt optical correlator, instead of using the reference beam to

reconstruct the stored signal, the signal beam is used to reconstruct the reference beam as

shown in Figure 4.1 (b). If the data pattern on the SLM is 'similar' to the stored hologram,

then the reference beam is reconstructed. The strength of the reconstructed reference

beam is proportional to the similarity between the input pattern and the stored pattern.

The reconstructed reference beam is then focused by a lens and a detector array reads its strength.

Since the comparison between the input image and the stored filter is done in the frequency plane (Fourier plane), the 'similarity' measurement is just the correlation function :

$$g(x', y') = \iint f(x, y)h(x - x', y - y')dxdy \qquad (4.1)$$

where $f(x,y)$ and $h(x,y)$ are the input image and the stored filter, respectively. (x,y) are the input coordinates at the SLM and (x',y') are the output coordinates at the detector array [4,5]. Eq. 4.1 is a valid expression for the Vander Lugt optical correlator if an infinitely thin recording material is used to store the filters. However, since planar recording materials usually have a very limited dynamic range, not many filters can be multiplexed at the same location. In order to have a large number of filters, we have to use a thicker recording material. One recording material that is nearly perfect for this application is DuPont's photopolymer. It is relatively thin (some versions are as thin as 15 μm) while exhibiting a relatively large dynamic range (approximately 50 times better than a piece of LiNbO$_3$ of the same thickness). The extra thickness of the filters modifies the optical correlation function to :

$$g(x', y') = \iint f(x, y)h(x - x', y - y')dxdy \cdot \text{sinc}\left(\frac{L\sin\theta_r}{\lambda f}x'\right) \qquad (4.2)$$

where L is the thickness of the recording material, θ_r is the angle of the reference beam with respect to the signal beam, λ is the wavelength and f is the focal length of the Fourier transforming lens. The extra sinc() term in Eq. 4.2 is due to the volume hologram effect which reduces the shift invariance in the plane of interaction (the x direction in this case).

The normal correlation function is 2-dimensionally shift invariant, meaning if the input image is shifted in the x-y plane with respect to the pattern stored in the filter, the strength of the cross-correlation between them remains unchanged. However, the position of the cross-correlation peak will move in the x'-y' plane to indicate that the input image has shifted in position. For an optical correlator system using volume holograms as filters, if the input image is shifted in the x direction (Figure 4.2) then the angle at which its Fourier transform is incident on the recording material also changes. The angular change between the incident signal beam and the stored hologram causes a Bragg-mismatch effect very similar to angle multiplexing. As the input image is shifted in x, the intensity of the cross-correlation peak diminishes as a function of $sinc()^2$ while it moves along x'.
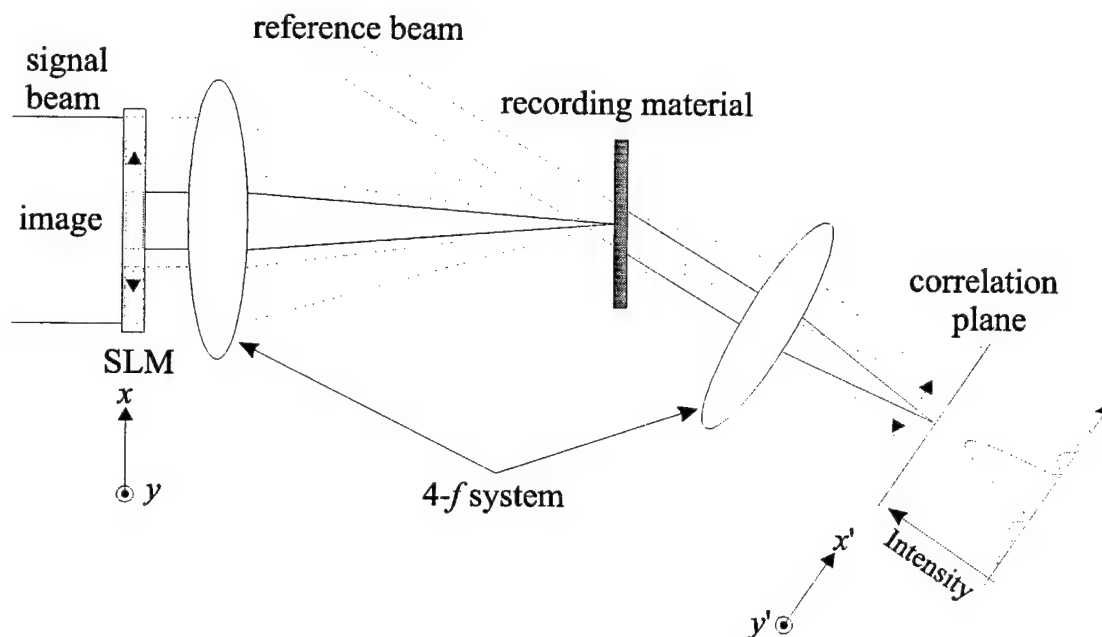

Figure 4.2 : Shift invariance in the plane of interaction for a volume holographic optical correlator.

With enough shift in the x direction, the cross-correlation peak can disappear altogether when it reaches the first null of the sinc() function at $\lambda f / L \sin \theta_r$. Therefore, in

order to have the greatest shift invariance (to better approximate the normal correlation function), we could increase the wavelength of the source and the focal length of the lens or decrease the thickness of the recording material and the angle of the reference beam. We could also use the smallest SLM possible to minimize the amount of displacement an image can have from the origin. For example, using λ = 532 nm, f = 50 cm, L = 38 μm, θ_r = 30°, an image can be shifted in the SLM by 1.4 cm in the x direction before reaching the first null. If the input image is shifted in y, the intensity of the cross-correlation peak remains largely unchanged since that is in the Bragg degenerate direction.

So far we have only talked about correlating the input pattern with one stored filter (hologram). By multiplexing multiple filters at the same location (using for example angle multiplexing), an input pattern can be cross-correlated with all the stored filters simultaneously. The reference beam for the best matched filter will be reconstructed the strongest. From the intensity of the correlation peaks, we can determine which filter has the best match. Furthermore, from the position of the brightest peak, we can tell if the input image is shifted with respect to the filter. This will become important later for automatic target tracking.

Many other types of optical correlators exists [6-11]. The joint-transform-correlator (JTC) [12] and the wavelength multiplexed - reflection geometry Vander Lugt correlators [13] are particularly interesting. Both types of correlators offer full 2-D shift invariance. The JTC has no pre-recorded filters, therefore the thickness of the filters is essentially zero. The reflection geometry Vander Lugt correlator's reference beam is in the counter-propagating direction with respect to the signal beam, therefore $\sin \theta_r$ = 0 and the sinc() function is 1 for all x'. The major drawback of these systems is that typically,

the input image can only be correlated with one filter at a time. The JTC requires a separate memory unit to store and update the filters while the reflection geometry Vander Lugt correlator requires a wavelength tunable laser source to access the stored filters one at a time. Being able to correlate against only one filter at a time seriously degrades the computational rate of these system. As far as I know, the Vander Lugt optical correlator setup shown in Figure 4.1 (b) still offers the best computational rate for a large database of filters, using currently available components.

For a rigorous derivation of the Vander Lugt optical correlator using volume holograms, I refer the readers to Sidney Li's thesis [5], pages 200 to 206.

4.2 Autonomous Vehicle Navigation

4.2.1 Introduction

A machine that uses visual sensing has access to large amounts of data about the surrounding environment. However, extracting useful information from this large data set is a difficult task given that serial computers can be easily overwhelmed by the sheer volume of data to be examined and by the amount of memory required. A single frame of uncompressed NTSC video typically contains over 250,000 bytes of data. There has been recent advances in digital-signal-processing (DSP) hardware and control algorithms that have allowed autonomous vehicles to recognize lanes, fork in road, and landmarks in real-time [14, 15]. In this chapter we present a memory rather than a model-based approach to navigation and target recognition/tracking using holographic data storage.

Image processing is well suited for optical implementation because of the inherent 2-D parallelism optics can provide. Furthermore, optical data storage offers high storage

capacity and fast parallel access to the stored information. If we combine optical image processing and optical data storage, we can have a system that is well suited to machine vision applications. Specifically, we have constructed a system using an optical correlator as the computational element and a holographic database as the storage device. With our system, an input image is simultaneously correlated with a large number of templates in the holographic database. Systems that combine holographic storage with correlation can produce up to 20,000 correlations per second (an input image updated at 30 Hz correlated against 1,000 stored templates). By measuring the position and intensity of these correlation peaks, a control algorithm can be devised to perform complex tasks.

In this section we describe an experimental vehicle navigation system that uses optical correlation to direct a vehicle along a predefined trajectory. In this experiment, we wanted the vehicle to be able to autonomously retrace a path that it had previously traveled under manual control. The vehicle uses only visual input to perform this task. The only sensor mounted on the vehicle is a CCD camera. Our approach to this problem is simple. A camera on the vehicle records the observed visual scene as the vehicle is manually driven along the desired trajectory. Sequential still-frames from this video are then selected to be 'waypoints' and they are stored in the holographic database. The vehicle retraces the desired path by simply driving from waypoint to waypoint until the final destination. The vehicle recognizes the waypoints through the correlation process. Furthermore, the correlation is done against many waypoints simultaneously so the vehicle can look ahead to upcoming waypoints to decide its future actions. The vehicle can also be trained to recognize and stalk targets by storing images of the target in the holographic database.

4.2.2 Experimental Setup

Figure 4.3 shows a photograph and the schematic diagram of the vehicle used in our experiment. The vehicle, which we constructed using off-the-shelf components, has three wheels and carries a miniature CCD camera, a video transmitter, a remote-control receiver, two drive motors, and two lead-acid batteries. It is 38 cm long, 31 cm wide and stands about 20 cm tall (not including the antenna mast). At full speed, the vehicle can travel ~10 cm per second forward or backward. Furthermore, since there are two independent drive wheels, the vehicle can turn and also spin on its own axis clockwise or counterclockwise.
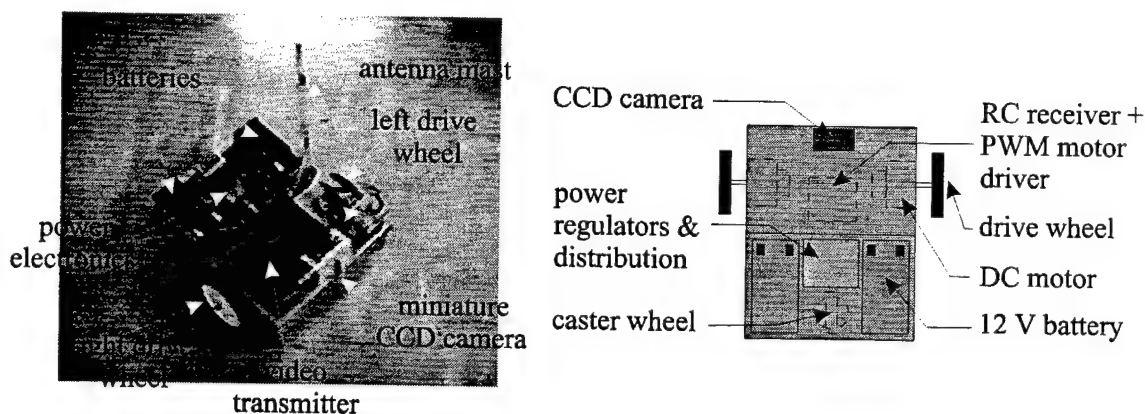

Figure 4.3 : A picture and schematic diagram of our vehicle.

The vehicle itself has no intelligence. It transmits the video from the onboard camera to the optical correlator and receives commands to drive the wheels. All the computation is done within a modified Vander Lugt optical correlator shown in Figure 4.4 and a personal computer. Real-time video is presented to the optical correlator through a spatial-light-modulator (SLM). The SLM used in this experiment was removed from an InFocus TVT-6000 video projection system. It has 480 x 440 pixels and a 3 cm

× 2.5 cm clear aperture. The SLM is illuminated with collimated light from a 200 mW doubled Nd:YAG laser ($\lambda = 532$ nm). In Figure 4.4, lens C2 is a 100 cm focal length lens and it is placed a few cm after the SLM. It produces the Fourier transform of the input image at plane C3. Such a long focal length lens was used to ensure good shift invariance of the correlation and to produce a large Fourier transform pattern at plane C3. A cross-hair spatial filter is placed at C3 to accomplish three goals. It blocks the higher diffracted orders that result from the pixelation of the SLM. Spherical aberration in lenses C2 and C4 limits the precision in which the higher diffracted orders can be collected. The removal of the higher orders gives a smoother image but it reduces the light efficiency of the SLM. The second function of the spatial filter is to block the low-frequency components of the input image. This edge-enhances the input image and greatly reduces the cross-correlation between different images. Finally, a thin cross-hair removes some of the horizontal and vertical lines from the input image in order to sharpen the correlation peaks. The vehicle navigates in the hallways of a building, where there are many vertical lines representing door ways and horizontal lines representing floor and ceiling tiles. If these horizontal and vertical lines are not attenuated from the input images, then the correlation peaks will also be in the shape of a cross and it will be difficult to determine the exact location of the peaks.

Lenses C4 and C6 image with magnification one plane C3 onto plane C7, the plane of the hologram. A second spatial filter is placed between the two lenses at the image plane of the SLM (considering lenses C2 and C4 as an imaging system). The purpose of this spatial filter is to remove the sharp rectangular edges of the SLM itself. These edges are common to all the images and they are removed to minimize cross-

correlation. The aperture of the filter is approximately 3 mm × 2.5 mm, since lenses C2

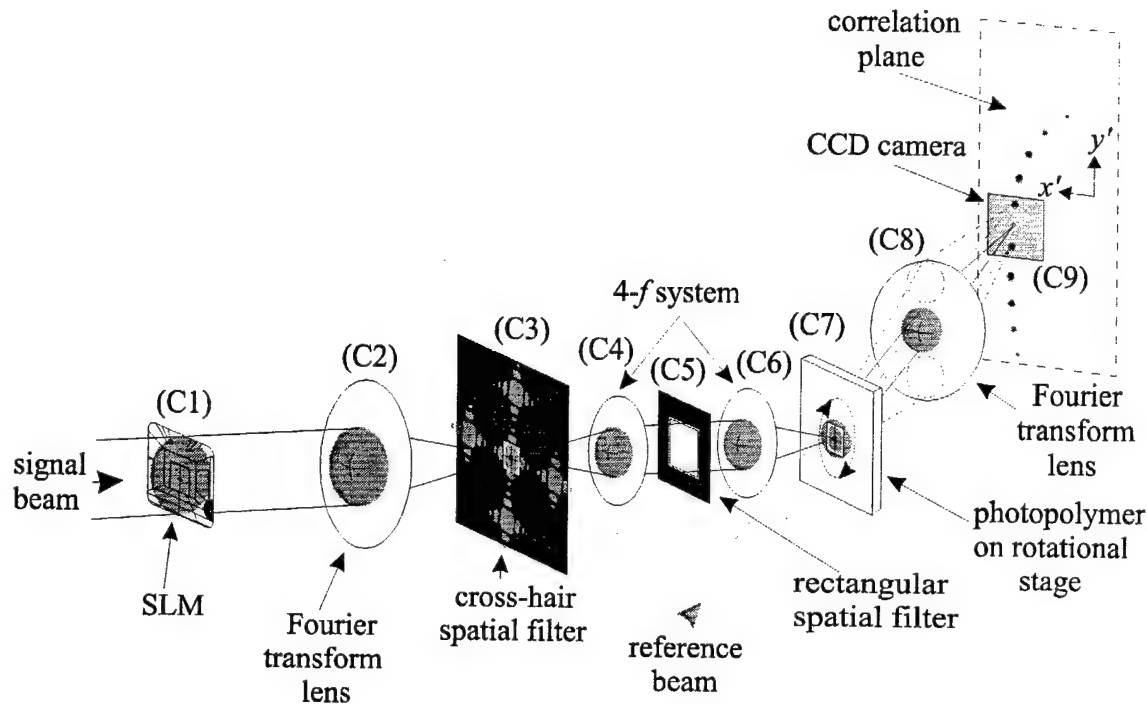and C4 form a 10:1 reduction system (the focal length of lenses C4 and C6 are 10 cm).

Figure 4.4 : A schematic diagram of the modified Vander Lugt optical correlator used in this experiment.

The filtered Fourier transform of the input video (at plane C3) is imaged onto the

recording material at plane C7. The recording material we use for this experiment is a

100 micron thick photopolymer film (DuPont's HRF-150-100). The DuPont

photopolymer is well suited for this experiment because the film develops itself without

having to remove it from the optical system for wet processing. This is nice since

alignment is fairly critical in an optical correlator. Furthermore, the photopolymer is a

write-once-read-many material. Once the recording process is complete, subsequent

exposure to light does not decay the previously stored images. Therefore, the correlation

process can be continued indefinitely without the need to refresh the stored holograms.

Finally, the DuPont photopolymer has very high diffraction efficiency for a thickness of

only 100 micron. We have found from previous experiments that the $M/\#$ (Chapter 2.4.4) of this material is approximately 6.5. A high $M/\#$ allows many holograms to be multiplexed on top of each other and still be recalled with sufficient diffraction efficiency. Furthermore, since the recording material is only 100 micron thick, we were able to make the volume hologram effect negligible by using a 100 cm focal length Fourier transform lens to obtain full 2-D shift invariance from the optical correlator. By contrast, in order to get comparable $M/\#$ from a $LiNbO_3$ crystal, we would have to use a thickness of approximately .5 to 1 cm, which would result in very limited shift invariance.

The photopolymer laminated to a 1 mm thick glass substrate is mounted on a rotational stage at plane C7. The photopolymer is placed normal to the signal beam and the center of rotation of the rotational stage is aligned with the optical axis of the signal beam. A planewave reference beam that forms an angle of 30 degrees (outside) with the signal beam is introduced to record holograms. Multiple holograms are stored through peristrophic multiplexing. After a hologram is recorded, the photopolymer is rotated in-plane by 1 degree before the next hologram is recorded. For a correlator, the peristrophic rotation moves the correlation peaks in the trajectory of a circle around the signal beam (see Chapter 2.3). The 1 degree clockwise peristrophic rotation moves the correlation peak of the previously stored hologram slightly above the focal point of the reference beam as shown in plane C9 of Figure 4.4. This allows another filter to be stored at the same location on the photopolymer and still have spatially distinct correlation peaks. This process continues until all the filters have been stored (up to 180 filters using peristrophic multiplexing alone). Since the holograms are all stored in the same area, all the correlation peaks are produced simultaneously when an input image is presented.

Fourier transform lens C8 produces at the correlation plane the correlation peaks. Three of these peaks are monitored by a CCD camera, C9, placed at the correlation plane. Other correlation peaks can be brought into the field-of-view (FOV) of the CCD camera by rotating the photopolymer film.

To determine the proper action for the vehicle to take, we decipher the video from the CCD camera placed at the correlation plane. The information of interest is the intensity and the position of the peaks. We will ignore and suppress other information such as the structure of the correlation peaks, which contains clues about the environment the vehicle is in. We used a personal computer to extract the intensity and the position of the peaks. The video from camera C9 is fed into the computer's memory through a framegrabber board. A search algorithm locates and determines the intensity of the three correlation peaks. To follow the waypoints stored in the holographic database, the computer issues commands according to the flowchart shown in Figure 4.5. First, the computer locates the brightest peak (the CCD pixel array is turned into a Cartesian coordinates system with the origin at the center of the CCD array). If the brightest peak is centered in x' within a pre-assigned margin ε, then the vehicle is assumed to be on path and both wheels are set to their nominal velocity to drive straight ahead. A centered peak indicates that the vehicle's viewpoint is the same as when the vehicle was driven manually down the path the first time (see Figure 4.6). However, if the peak position is shifted to the left or right, the current viewpoint of the vehicle is shifted with respect to the stored waypoint. This happens when the vehicle is heading in a different direction than the original direction. The computer then issues commands to drive one of the wheels faster to turn the car back to the desired heading.
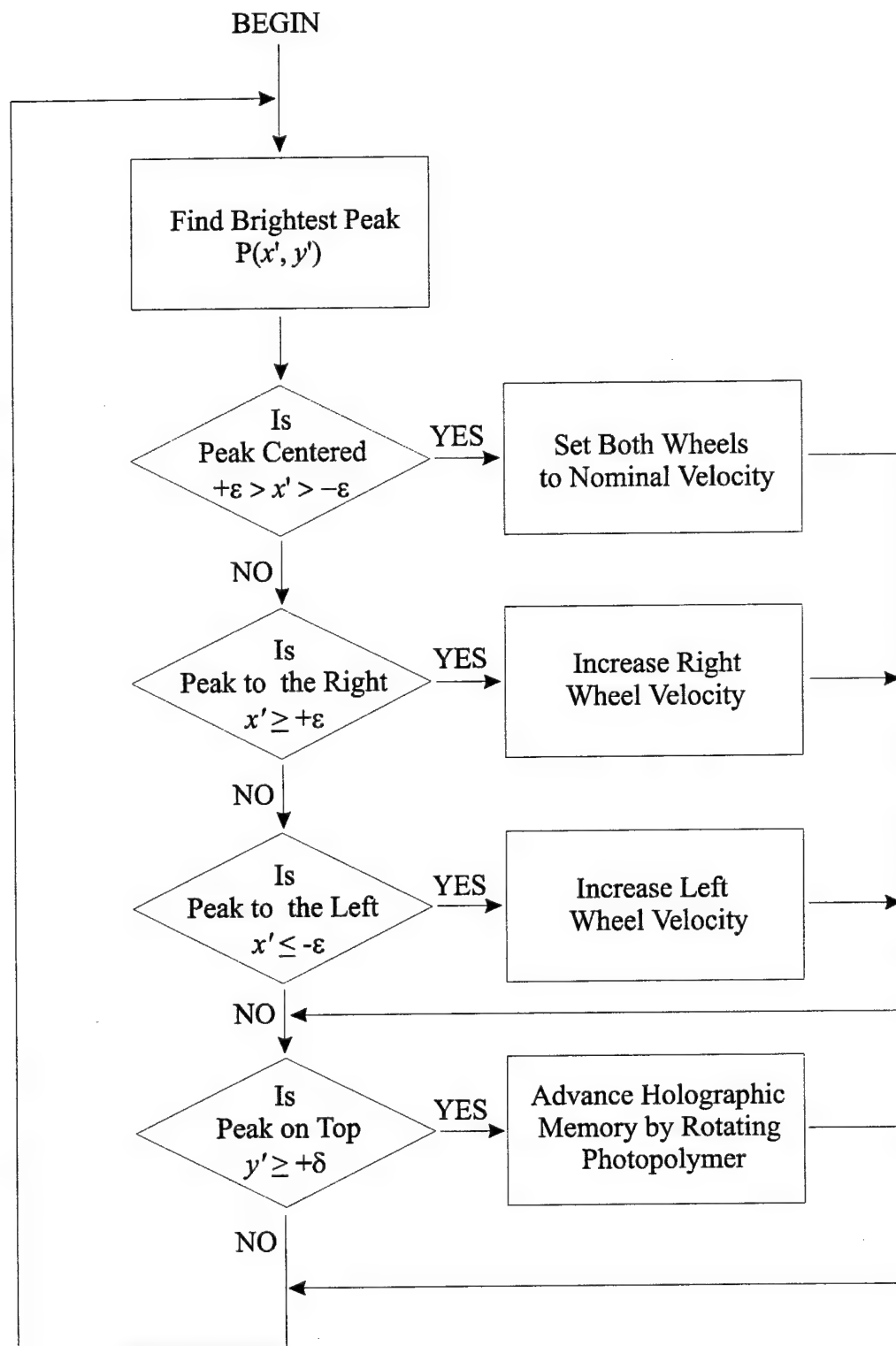
Figure 4.5 : The flowchart of the basic vehicle control algorithm in the computer.

We monitor 3 peaks with CCD array, C9, placed at the correlation plane. The brightest peak is suppose to be in the middle of the CCD array and it represents our current position on the path (see Figure 4.7). The top peak represents the next waypoint the vehicle is looking for and the bottom peak represents the last waypoint the vehicle was at. In this way, we can keep an eye on what is coming up next and where we were before. As the vehicle progresses down the path, eventually the top correlation peak will become brighter than the middle peak. This means that the vehicle has reached the next waypoint and the computer issues a command to advance the holographic memory by rotating the photopolymer film to bring the top peak to the middle. Now a new waypoint appears on top and the old last waypoint drops out from the FOV of the CCD array. The above process is repeated continuously until the final waypoint has been reached and thus the path is complete.
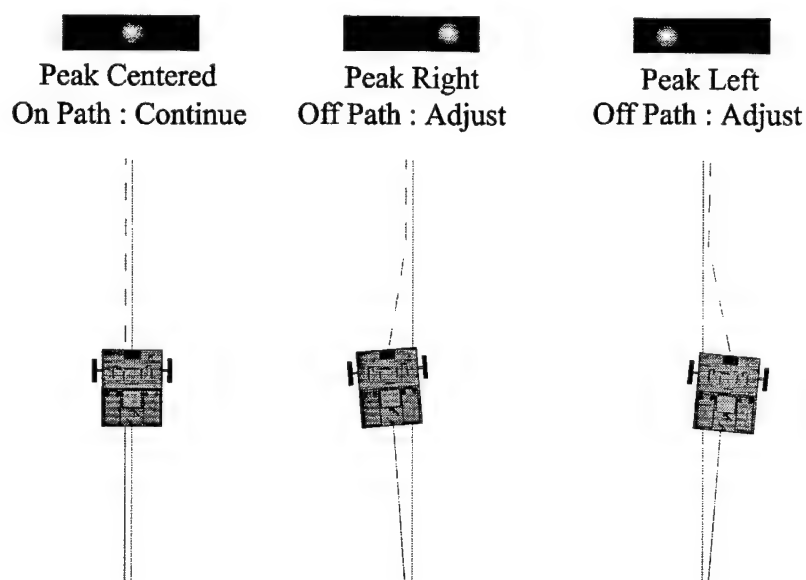

Peak Centered
On Path : Continue

Peak Right
Off Path : Adjust

Peak Left
Off Path : Adjust

Figure 4.6 : The position of the brightest peak in x' determines if the vehicle is on path.

Figure 4.8 shows an overview of the entire system. The vehicle sends the live video back to the optical correlator. The system computes the current position and the

trajectory of the vehicle from the correlations with stored waypoints and the location of the brightest peak. It sends control signals to the vehicle, and advances the holographic memory. Notice that once the control algorithm has been finalized, the computer and the CCD array at the correlation plane can both be replaced by a single VLSI chip. Since the control algorithm is just a simple set of hardwired decisions, a VLSI chip with CMOS detectors and simple logic would make the optical correlator a self-contained system with input (video from vehicle), computation (correlation), memory (holographic database), and output (commands to drive the vehicle).
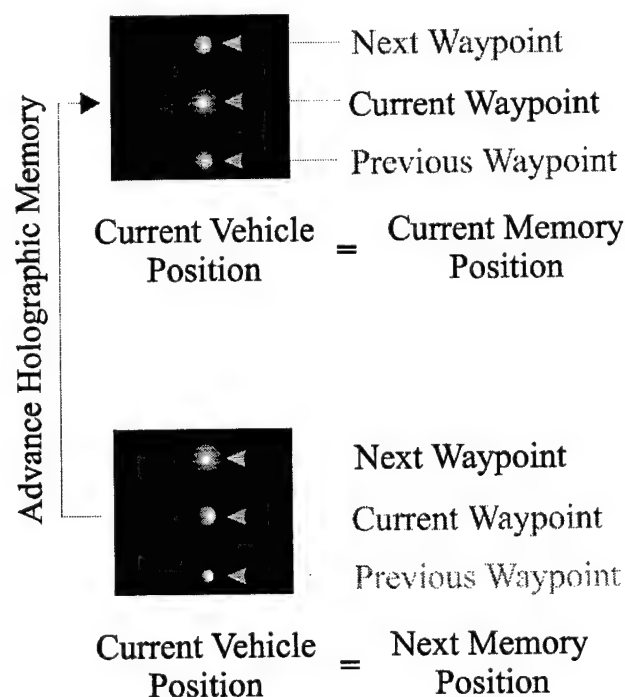

Figure 4.7 : The position of the brightest peak in y' determines if the vehicle has reached the next waypoint.
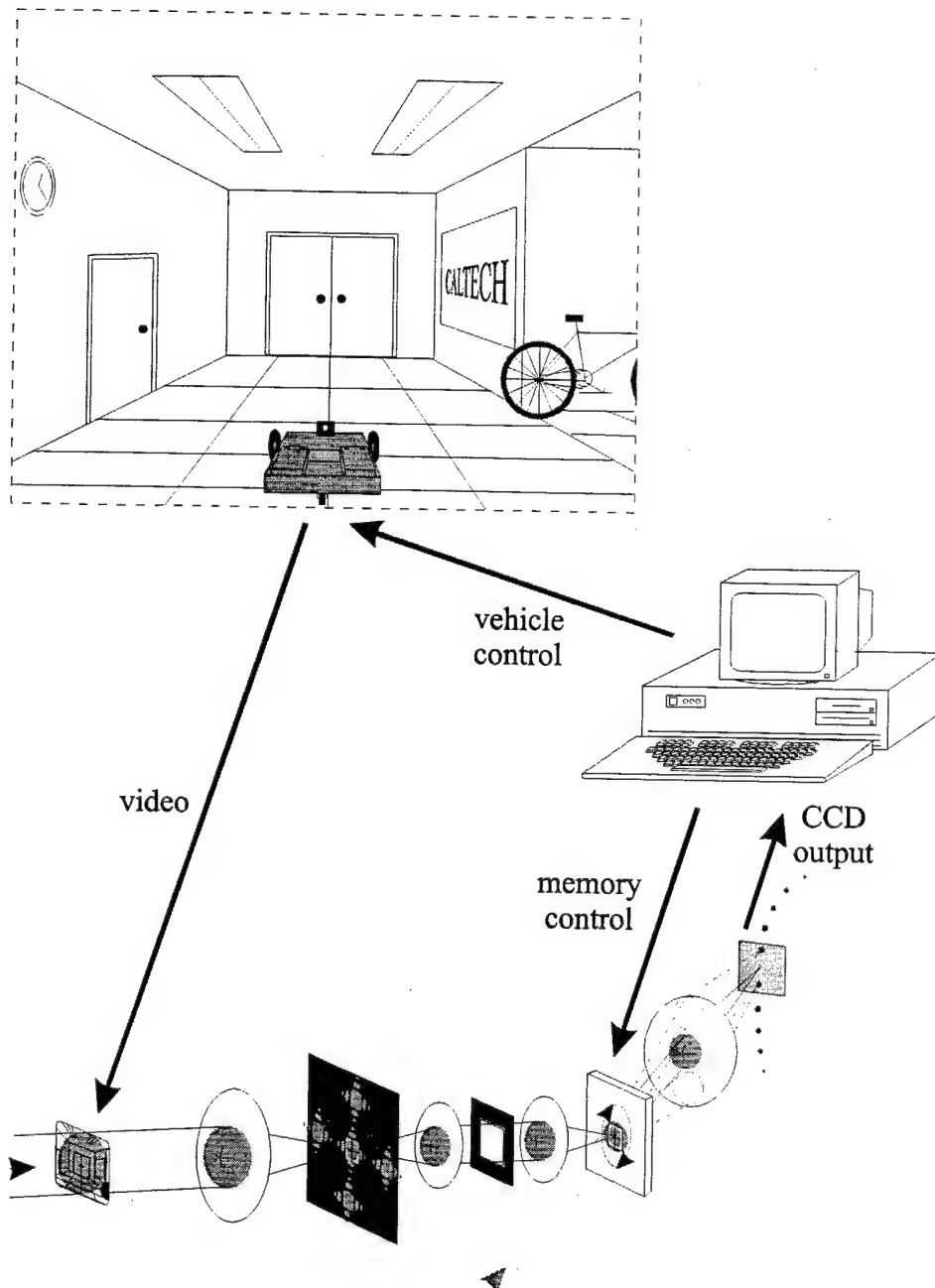
Figure 4.8 : An overview of the entire autonomous vehicle navigation system.

4.2.3 Modes of Navigation

The flowchart shown in Figure 4.5 describes how the vehicle follows the stored waypoints to the final destination. We call this the 'follow mode'. However, if the

vehicle is pointed in a direction so that it sees an unfamiliar scene, then no correlation peak is produced and the vehicle would be lost. In this case, the vehicle switches to the 'acquisition mode' where it rotates in place. More specifically, if the top correlation peak is undetectable, then the vehicle is instructed to rotate in place until the top peak becomes visible again. This is a logical choice since if the vehicle cannot see the next waypoint, there is no point in going ahead and getting further lost. The acquisition mode is also useful when the vehicle's field-of-view is blocked temporarily by an object. When that occurs, the top correlation peak goes blank and the vehicle would simply rotate in place until the interfering object is removed.

With follow and acquisition modes, we can navigate the vehicle through the desired path even when the vehicle experiences unexpected perturbations. To further demonstrate the utility of our system, we decided to have the vehicle find and then stalk a target once the vehicle has reached a destination. The 'stalking mode' is just a simple addition to the control algorithm and it works as follows: A far and near image of a target is stored after the last waypoint in the holographic memory (far and near correspond to the distance away from the miniature CCD camera on the vehicle, see Figure 4.9). After the vehicle has reached the final waypoint, the computer advances the holographic memory to bring the far and near correlation peaks onto the CCD C9. If the target is present and is at the desired distance away from the vehicle, then the two correlation peaks will have equal intensity. Furthermore, if the correlation peaks are also centered with respect to x, then the vehicle is correctly positioned and will stop both wheels and fixate on the target. However, if one of the peaks is brighter than the other, then the vehicle will either backup or drive forward until both correlation peaks are of equal

intensity, which indicates that the vehicle is at the desired distance away from the target. If the peaks are not centered, then the vehicle will turn in place to bring the target back to the center of its FOV. If there are no detectable correlation peaks, then the target is not present in the current FOV and the vehicle is instructed to rotate in place until the target is found.
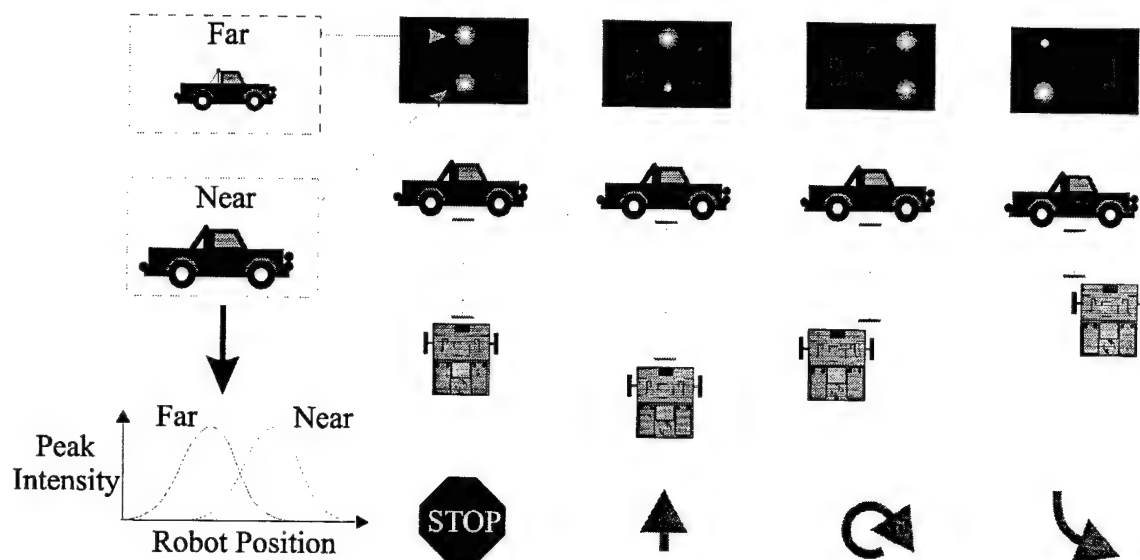

Figure 4.9 : Stalking mode operation.

Using these navigational modes, we can drive a vehicle from place to place and then find/stalk a target of interest. Other tasks can be accomplished by using a combination of these basic modes.

4.2.4 Holographic Database

Figure 4.10 shows a picture and the layout of the hallway the vehicle navigates. The width of the hallway is approximately 2.4 meters and the floor is covered with glossy square tiles (23 cm x 23 cm per tile). The video output from the onboard miniature CCD camera was recorded on a laser disc as the vehicle was driven manually through Path 1

from Room A to B. We take still-frames from this video to store in the holographic database as waypoints for the vehicle to follow. A video sequence was also recorded for a second path. Path 2 starts from Room C and goes up to the intersection point with Path 1. With waypoints extracted from these videos, the vehicle can start in anyone of the Rooms and navigate to any other.
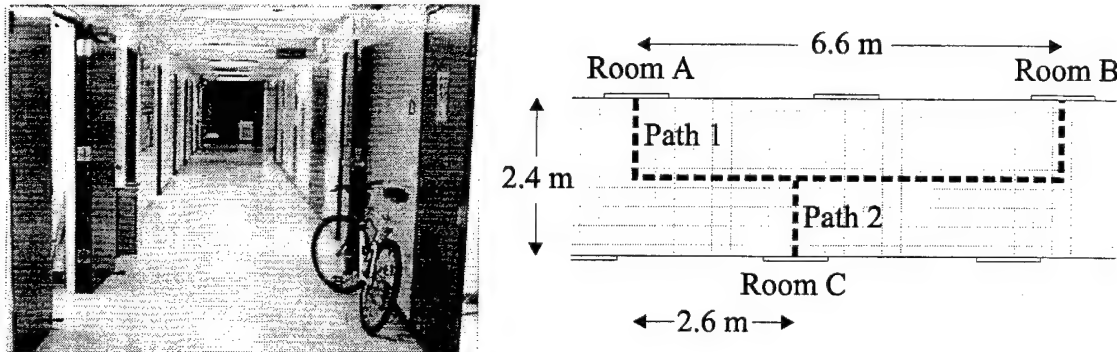

Figure 4.10 : A picture and the layout of the hallway the vehicle navigated.

How far apart should the waypoints be separated (in hallway distance) for our system to function properly? If the waypoints are separated too far, then the vehicle will be lost since the top correlation peak corresponding to the next waypoint will be dim and undetectable by the time the middle correlation peak fades away. On the other hand, if the waypoints are too close together, then we will be over-sampling the path and wasting memory on redundant information. Initially we selected the waypoint separation so that the top correlation peak is slightly above the detectable threshold when the middle correlation peak reaches maximum intensity. This way, the vehicle will always be able to see the next waypoint as it navigates the path. However, since the hallway is sometimes noisy (pedestrians or foreign objects blocking parts of the hallway), we found that it is better to ease off from this separation criterion and have some redundancy by selecting waypoints that are closer in distance.

Through trial-and-error, we found that waypoints separated by one floor tile (23 cm) gives the desired performance in our environment. Of course the required separation would depend on the specific environment. For example, if the hallway was 1 meter in width instead of 2.4 meters, then the waypoints might have to be closer together. This is due to the fact that the rate of change of the surrounding as seen by the moving vehicle is greater in the 1 meter hallway than with the 2.4 meter hallway. Similarly, for a Boeing 747 cruising at 35,000 feet, the rate of change of the earth below is slow and a new waypoint might not be needed for tens of miles.

Figure 4.11 shows the 46 waypoints needed to describe Path 1 from Room A to B (order of the sequence is from left to right, in a raster scan format). The first 9 frames show the vehicle driving out of Room A and approaching a parked bicycle. The bicycle was placed in front of Room A to add structure to the wall. Without the bicycle, the vehicle would be looking at a nearly blank wall and the correlation function will fail. Frames 10 through 38 define the straight path down the middle of the hallway. Fortunately, the hallway itself is full of bulletin boards and door frames so that additional structure is not required. The sharp transition from frame 9 to 10 is intentional so that the vehicle would get lost temporarily. When the vehicle reaches the location of frame 9, the next waypoint the vehicle is looking for is frame 10 (down the hallway). Since the vehicle is still looking at the bicycle, the top correlation peak corresponding from frame 10 is undetectable. Using the control algorithm discussed above, the vehicle is instructed to rotate in place until the next waypoint is found. After the vehicle has rotated far enough so that it is looking down the hallway, the top correlation peak will become bright enough to be detected and the holographic memory is advanced to bring the top peak to

the middle. The vehicle then continues navigating on the follow mode. We use the acquisition mode to make turns in order to save on memory. The alternative would be to store the turn itself as a sequence of waypoints and use the follow mode to complete the turn. This would add an additional 4 to 6 waypoints for each 90 degree turn, or a 20% increase in memory requirement for Path 1. The final eight frames of Figure 4.11 shows the vehicle driving toward Room B. Once the vehicle reaches the last waypoint, it is instructed to stop. Path 2 is described by the five waypoints shown in Figure 4.12. When the vehicle reaches the last waypoint on Path 2, it turns (using acquisition mode) and follows frame 20 of Path 1 to reach Room B.

The combined total of 51 waypoints (46 from Path 1 and 5 from Path 2) are stored in the DuPont HRF-150, 100 micron thick photopolymer using peristrophic multiplexing. Three correlation peaks fit on the active area of the CCD array C9 given the 1 degree peristrophic rotation and the focal length of the lens C8. The holograms were recorded with a signal and reference beam intensities of approximately 500 μW/cm^2 each. It is desirable for the maximum correlation peak intensity for all the waypoints to be the same or the control algorithm might not advance the holographic memory properly during the experiment. The peaks can be equalized by simply varying the recording time of the waypoints according to the number of edges it contains. For waypoints with many edges, a shorter recording time would produce a correlation peak that is of the same intensity as a waypoint with fewer edges recorded for a longer duration. For Path 1, the recording time varied from 1 to 3 seconds per waypoint and the recording schedule was obtained through trial-and-error. At first, all the waypoints were stored with the same exposure time. The waypoints were then played back from the laser disc into the optical correlator

for auto-correlation. For noticeably stronger or weaker auto-correlation peaks, their recording time was decreased or increased accordingly. A recording schedule that produced visually equalized correlation peaks was obtained after 4 trials. For the 51 waypoints that were recorded, we stayed well within the linear region of the photopolymer and did not approach dynamic range saturation.

The 1 degree peristrophic rotation between holograms has another interesting and desirable effect. The hologram representing the top correlation peak is rotated by +1 degree (in orientation) with respect to the hologram representing the middle peak (the middle hologram is at 0 degree rotation with respect to the input SLM). Similarly, the hologram representing the bottom correlation peak is rotated by -1 degree with respect to the middle hologram. During correlation, this ± 1 degree of rotation diminishes the peak intensity of the top and bottom correlation peaks while leaving the middle peak unaffected. This is due to the fact that the correlation function is not rotational invariant. What this means is that the vehicle has to drive close to the next waypoint before the top correlation peak becomes brighter than the middle peak. This inhibition of the top (and bottom) correlation peak is desirable because it prevents the computer from advancing the holographic memory too far ahead of the vehicle's current position, which could occur if not all the waypoints have similar maximum peak intensity.
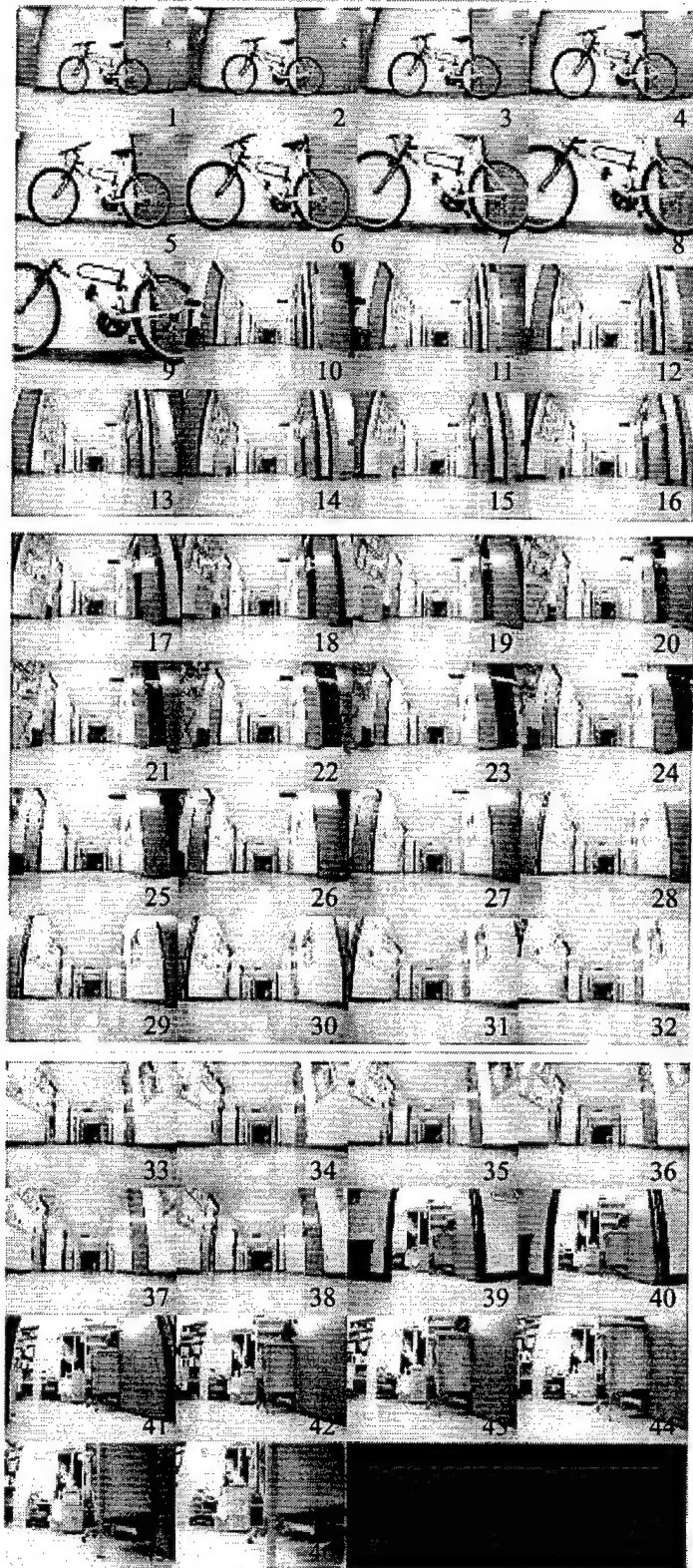
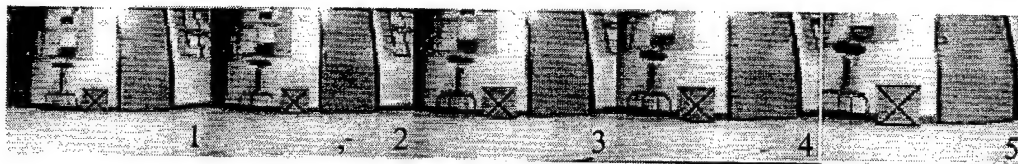Figure 4.11 : The 46 waypoints stored to describe Path 1.

Figure 4.12 : The 5 waypoints stored to describe Path 2.

4.2.5 Experimental Results

After the waypoints have been properly selected and recorded, the vehicle was able to consistently and accurately retrace the programmed path using a combination of follow and acquisition modes. Figure 4.13 shows a sequence of the vehicle navigating Path 1. Each frame consists of the video from the miniature CCD camera (main view), the correlation peaks at the correlation plane (upper left-hand corner), and the image of the waypoint corresponding to the top correlation peak (upper right-hand corner). When the vehicle first starts from Room A (frame 1 of Figure 4.13), it only has two correlation peaks corresponding to the current position and the next waypoint. It doesn't have a previous position stored in its memory so the bottom correlation peak is blank. As the vehicle moves forward toward the bicycle, the holographic database advances and all three peaks becomes detectable. Frame 4 of Figure 4.13 shows that the vehicle has reached the mid-point of the hallway and it should execute a turn to the port side at this point. Notice that the top correlation peak in frame 4 is undetectable since the next waypoint the vehicle is looking for is completely different from its current view. This activates the acquisition mode and the vehicle begins to rotate in place. Frame 5 shows the view from the vehicle in the middle of the rotation and all the correlation peaks have gone blank. After the vehicle has rotated far enough, its viewpoint is the same as the next

waypoint so the top correlation peak becomes detectable again (frame 6). The control algorithm switches from acquisition mode back to follow mode and advances the holographic memory to bring the top peak to the middle (since the top peak is brighter, frame 7). Frame 8 shows the vehicle navigating using follow mode as it had before in frames 1, 2 and 3. In follow mode, the velocity of the vehicle was 7 cm per second for the experiment shown in Figure 4.13. Later, with an improved control algorithm, the velocity approached the limit of vehicle at 10 cm per second.
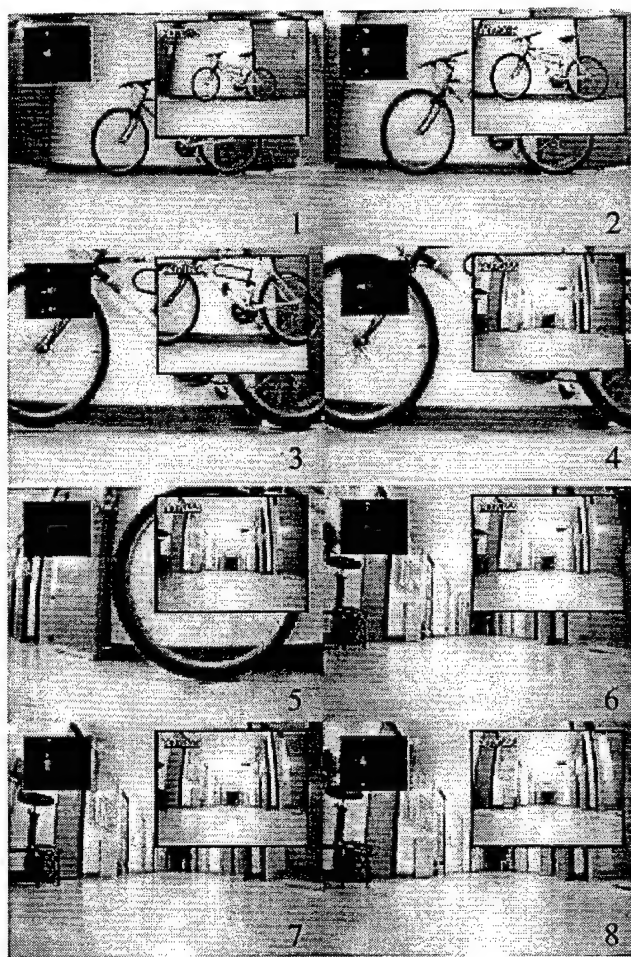
4:13 : Sequence taken from an experiment where the vehicle navigated Path 1.

The computer monitoring the correlation peaks is aware of the position of the vehicle with respect to the hallway. The computer knows which correlation peak is

currently the brightest and by reading the angle on the rotational stage, it can pinpoint the current waypoint the vehicle is at. This information is useful for more complex instructions such as navigating the vehicle to a particular waypoint and then branch off to join another path. For the experiment where the vehicle was navigated from Room C to Room B, the computer kept track of the vehicle's position on Path 2. After the vehicle reached the fifth and last waypoint of Path 2, the computer instructed the holographic memory to scan quickly to place frame 20 of Path 1 as the next waypoint. The vehicle then used acquisition mode to find Path 1 and followed the waypoints to Room B. We can also program the vehicle to drive backwards from Room B to Room C or A by using the bottom correlation peak as the next waypoint and the top peak as the previous waypoint.

The sequence of frames in Figure 4.14 show the types of noise we introduced into the hallway to try to fool the vehicle. In frame 1, a table and a man doing jumping-jacks at the middle and the end of the hallway did not hamper the normal operation of the vehicle. A remote-controlled truck ramming the vehicle as shown in frame 2 was also unable to make the vehicle lose its way. The vehicle was deviated temporarily from its course but always recovered and was able to drive back onto the path. In frame 3, the vehicle was rotated manually while it was in the follow mode. This caused the correlation peaks to disappear and when the vehicle was replaced on the floor it rotated in place until it was once again aligned with the path and then continued to its final destination. Frame 4 shows that with up to 40%-50% of its FOV blocked, the vehicle was able to continue driving past the obstacle. However, had the person stepped directly

in front of the path of the vehicle, its FOV would be blocked sufficiently to cause the vehicle to rotate in-place until the obstacle is removed.
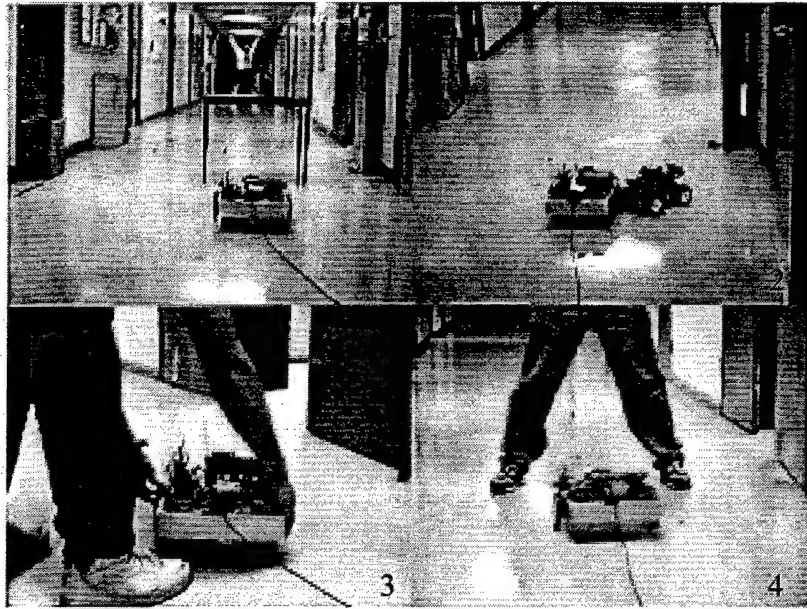

Figure 4.14 : Types of noise we introduced to test the performance of the system.

One way we could confuse the vehicle is to move it laterally away from the middle of the hallway by a few feet (toward one of the walls). This changes the perspective of the hallway as seen by the vehicle and correlation with the stored waypoints will fail. The vehicle becomes lost and rotates in-place without ever recovering the path. For normal operations, this scenario should rarely occur.

To demonstrate the stalking mode, we programmed the vehicle for a different path, from Room A into another room with a cleared center. At the end of the sequence of waypoints, a far and near image of a red monster truck was recorded and designated as the target. During the experiment, the vehicle navigated from Room A into the cleared room using the stored waypoints. Once the vehicle has reached the last waypoint, the computer switches into the stalking mode and rotates the vehicle in-place until the target

is in view. The sequence in Figure 4.15 shows the stalking mode in action. In frame 1, the vehicle has acquired the target located at the desired distance from the vehicle. Frame 2 shows the target moving towards the starboard side of the vehicle and the vehicle rotates in-place to track the target. If the target moves away from the vehicle, the vehicle drives forward until the desired distance has been reached (frame 3). Finally, if the target gets too close to the vehicle, the vehicle backs up and re-distance itself from the target (frame 4). So by using just two correlation peaks, the vehicle was able to stalk a moving target in real-time.
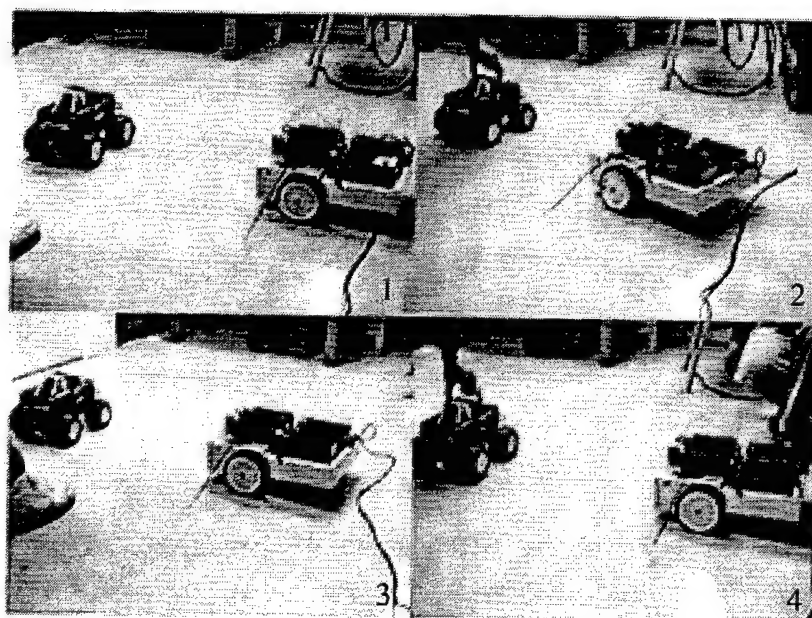

Figure 4.15 : Sequence showing the vehicle stalking a monster truck.

There are many ways in which the performance of this system can be improved. For navigation, instead of recording only a single path from Room A to B, down the middle of the hallway, several paths can be recorded. This will allow the vehicle to navigate around obstacles blocking a particular path. For stalking, images of the target at many different distances from the vehicle can be stored to extend the range of target

acquisition. More clever use of the correlation peaks, improvements in the control algorithm, and optimized correlation templates can further increase the capability of our system beyond its current simplistic form.

4.3 Automatic Target Recognition and Tracking

4.3.1 Introduction

The ability of optical correlators to process high resolution images in real-time can be applied to automatically recognize and track pre-programmed targets. The target is acquired by correlating the input scene with stored filters and then tracked by the 2-D shift invariant feature of correlators. However, the correlation function is not rotational or scale invariant and therefore if the scale or the prospective of the target changes, the target will no longer be recognized. Research in this field has generated several different approaches to recognize the target regardless of rotational and scale variations. One is to use feature based detection schemes [16], where a probabilistic model of the mutual positions of salient features can be used to recognize the target. This method can robustly recognize the target even in cluttered backgrounds but the generation of the probabilistic model for a complex target is sometimes difficult. Another method is to design a singular composite filter that incorporates multiples images of the target taken from different scales and rotations [17]. Since these composite filters have to recognize the same target regardless of orientation, they usually don't discriminate against anti-targets very well. It is possible to train composite filters with both targets and anti-targets to increase discrimination [18], but it is generally not possible to take into account all possible anti-targets that might be in the input scene.

Our approach to the scale and rotational sensitivity of optical correlators is simple. Since the holographic database can be very large, we will store all the 'interesting' views of the target in the database. Each filter is just one view of the target so the system can maintain good discrimination against anti-target and also determine the orientation of the target to within the difference between the filters. In this section several automatic target recognition and tracking experiments using this approach will be shown.

4.3.2 Target - Monster Truck

In the autonomous vehicle navigation experiment, the vehicle we constructed was an awkward platform for stalking targets. For example, to track the lateral motion of a target, the vehicle has to rotate its entire body with all the equipment onboard. The momentum associated with all this weight slowed down the reaction time of the vehicle. Therefore, we constructed a second vehicle based on a 1/16th scale model of the German Panzerkampfwagen VI - Tiger II tank (made by TAMIYA) for tracking targets. Furthermore in the autonomous vehicle navigation experiment, the system can only lock onto a single prospective of the target. In this experiment, we will store many different views of the target (each in two different scales) in the holographic database so the target can be tracked regardless of prospective and scale.

4.3.2.1 Experimental Setup and Control Algorithm

Figure 4.16 shows a picture of the Tiger II tank we modified for this experiment. This tank has a turret which can be rotated independently from the rest of the tank. We made extensive drive-train modifications and replaced the original turret with a small

aluminum box. The aluminum box houses a stripped-down SONY 8 mm camcorder rewired to allow computer control of its zoom functions. The tank and the camcorder are controlled by a computer through a digital wireless link. The range of motion of the tank is similar to the previous vehicle except now the tank can rotate its turret and inside the turret, the camera can be zoomed in or out. Therefore, a moving target can be tracked by rotating the turret and adjusting the zoom lens (instead of moving the entire tank).
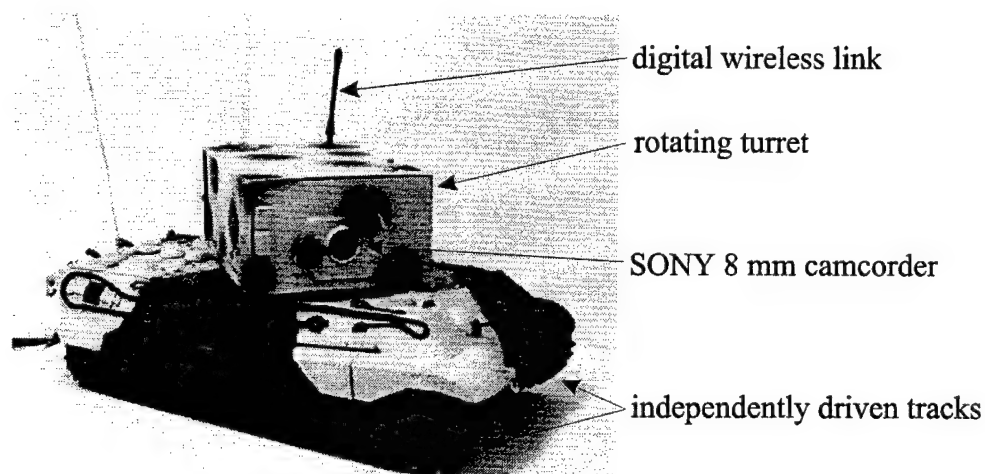

Figure 4.16 : A picture of the Tiger II tank we modified for this experiment.

The video from the camcorder inside the turret is transmitted via a wire to the optical correlator system shown in Figure 4.17. The optical correlator setup used in this experiment is similar to the one shown in Figure 4.4, except: (1) the SLM is mounted sideways so as the target moves laterally in the scene, it moves vertically in the optical correlator system (along y); (2) the cross-hair spatial filter is replaced with a simple DC filter; (3) a beam deflection system is placed in the reference arm and imaged to the recording material with a 4-f lens system; (4) 6 correlation peaks are monitored simultaneously with the CCD array placed at the correlation plane instead of 3.
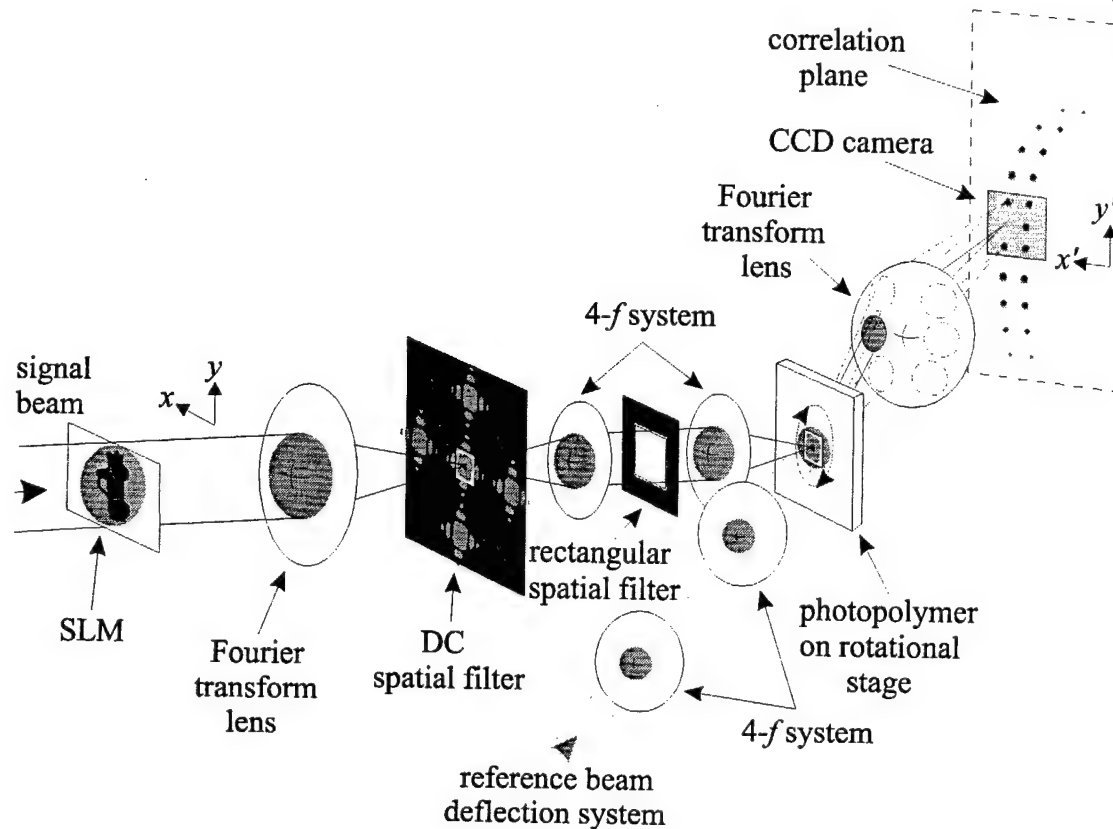
Figure 4.17 : The schematic diagram of the optical correlator setup used in the monster truck experiment.

Since the camcorder inside the turret is level with the target, most of the target's movement will be in the horizontal plane of the scene. By mounting the SLM sideways, we can exploit the Bragg degeneracy along the y-axis to increase the shift invariance. In the autonomous vehicle navigation experiment, a cross-hair spatial filter was placed in the Fourier plane of the SLM to block the numerous horizontal and vertical edges of the hallway. For automatic target recognition and tracking, it is desirable to keep all the edges of the target for correlation and to block the DC of the Fourier transform to improve discrimination. The DC filter we used was just a small (~.5 mm × ~.5 mm) piece of aluminum foil glued to a clear glass substrate. For this experiment, angle multiplexing was combined with peristrophic multiplexing by adding a beam deflection

system in the reference arm. A mirror mounted on a computer controlled rotational stage deflects the reference beam along the plane of interaction (along x), and the 4-f system images the deflected beam onto the recording material. When holograms are recorded with both angle and peristrophic multiplexing, the correlation peaks at the correlation plane show up as concentric circles, with the signal beam at the center. By selecting an angular and peristrophic separation of $1°$ between holograms, we can monitor 6 peaks simultaneously with the CCD array at the correlation plane (as shown in Figure 4.17).

To control the zoom lens of the camcorder inside the turret, the rotation of the turret, and the holographic database, we need to determine the intensity and the position of the correlation peaks. The control algorithm can be very simple if the holographic database (correlation peaks) is organized appropriately. In the autonomous vehicle navigation experiment, the waypoints were sequential and continuous. Therefore, peristrophic multiplexing was used to store the waypoints so they could later be accessed in the same order. This greatly reduced the number of correlation peaks that had to be monitored simultaneously to navigate the vehicle. The same can be done for tracking a target from many different perspectives. The holographic database of the target may consist of sequential but continuous perspectives of the target taken from many different angles. For each perspective, two or more scales of the target might be required to acquire the target from different range. To minimize the number of correlation peaks that the system has to monitor simultaneously, the holographic database (correlation peaks) can be organized as follows. Different perspectives of the target (at one scale) are stored peristrophically so they could be accessed sequentially in a continuous fashion. Different scales of the target are stored at separate angle multiplexing positions. At the correlation

plane, this creates concentric rings of correlation peaks, each ring represents the target at one scale from different perspectives. This is shown in Figure 4.18 for a target (monster truck) taken at 2 scales from many different perspectives.
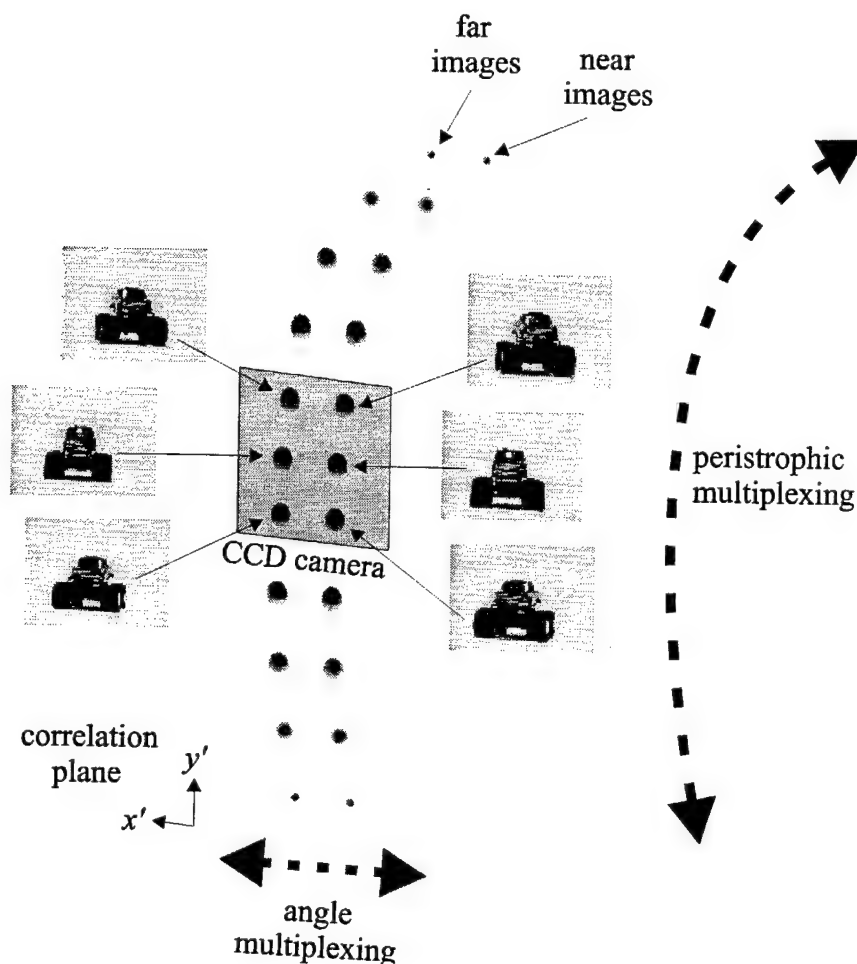

Figure 4.18 : The organization of the holographic database to minimize the number of correlation peaks that the system has to monitor simultaneously.

As stated previously, the CCD array at the correlation plane can monitor up to 6 correlation peaks simultaneously when the holograms are separated by 1° in both angle and peristrophic multiplexing. To generate the proper signals to control the zoom lens, turret, and holographic memory so the tank can track the monster truck regardless of position, orientation and scale, it is necessary to determine the intensity and position of

the peaks. A personal computer was used to extract those information and the computer issues commands according to the flowchart shown in Figure 4.19. Basically, the control algorithm brings the brightest peak to the middle of the CCD array (to track perspective and movement) and equalizes the far and near peaks to about the same intensity (to track scale). First the computer locates the brightest peak (the CCD array is turned into a Cartesian coordinate system with the origin at the center of the CCD array). If the brightest peak is in the upper section of the CCD array (one of the two top correlation peaks), then the computer advances the holographic memory to bring the upper section to the middle. When the peak is in the upper section, that means the perspective of the target is more similar to the holograms representing the top two peaks. By advancing the memory to bring the upper section to the middle, the next possible perspective of the target drops into the upper section and the previous perspective of the target moves to the lower section. Since the perspective of the target can only change in a continuous fashion as it is rotated, we can keep track of the target as it rotates by monitoring the next possible perspectives (peaks in the upper and lower sections). Similarly, if the brightest peak is in the lower section of the CCD array, then the computer rewinds the holographic memory to bring the lower section to the middle. To control the zoom lens of the camcorder inside the turret, we use the same differential detection scheme shown in Figure 4.9. If the brightest peak is in the right half of the CCD array, then the target is getting closer (the scale of the target is larger) and the zoom lens is instructed to zoom out. If the brightest peak is in the left half of the CCD array, then the target is moving further away and the zoom lens is instructed to zoom in. The only time the zoom lens is stationary is when the peaks in the middle section have similar intensity.
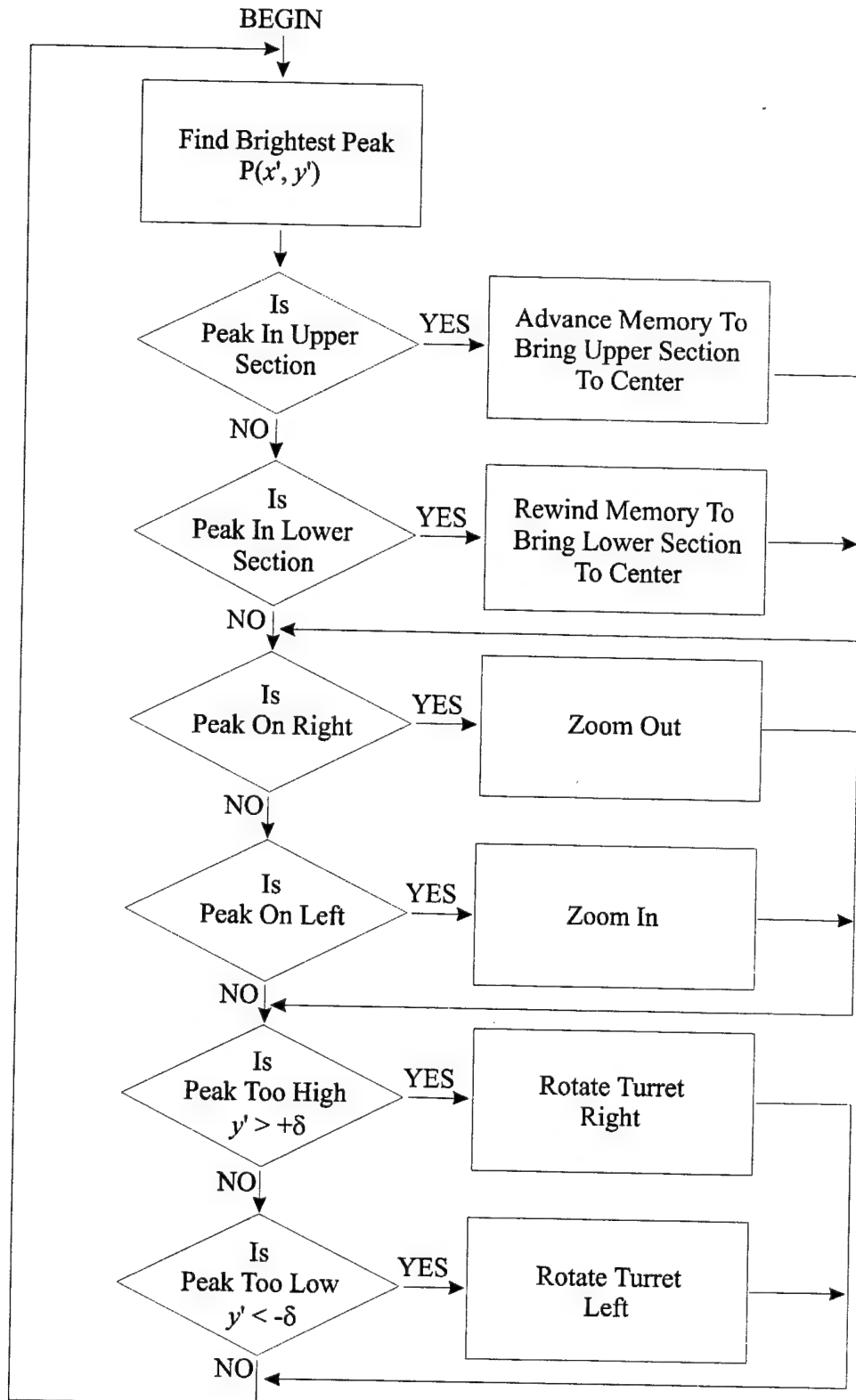
Figure 4.19 : The flowchart of the control algorithm for the tank.

The position of the correlation peaks in the middle section is used to rotate the turret. As the target moves laterally in the input scene, the correlation peaks would move vertically along y' since the SLM is rotated 90° to increase shift invariance. If the correlation peaks moves out of the dead-zone ($-\delta > y' > +\delta$), then the turret is instructed to rotate in the direction that brings the peaks back into the dead-zone. Using the control algorithm shown in Figure 4.19, the tank can keep the monster truck at the center of its FOV and always at the same scale, regardless of position, orientation and scale changes.

Figure 4.20 shows an overview of the entire system. The video from the camcorder inside the turret of the tank is sent to the optical correlator via a cable. The optical correlator computes the cross-correlations and a computer determines the best match. Depending on the position of the brightest peak, the computer sends control signals to the zoom lens, turret, and holographic memory in order to track the target.

4.3.2.2 Holographic Database

For this experiment, we would like to track the monster truck if it poses a 'threat' to the tank. The tank is threatened if the monster truck is pointed toward the tank. Figure 4.21 shows 19 different perspectives of the monster truck (as seen by the camcorder inside the turret of the tank) we consider a threat to the tank. For each perspective, a far and near image were taken to define the scale range. Consecutive perspectives of the monster truck were separated by 10° in rotation. This separation was experimentally verified to be fine enough in resolution so as the perspective of the monster truck changes, there is always a bright peak at the correlation plane to track the target.
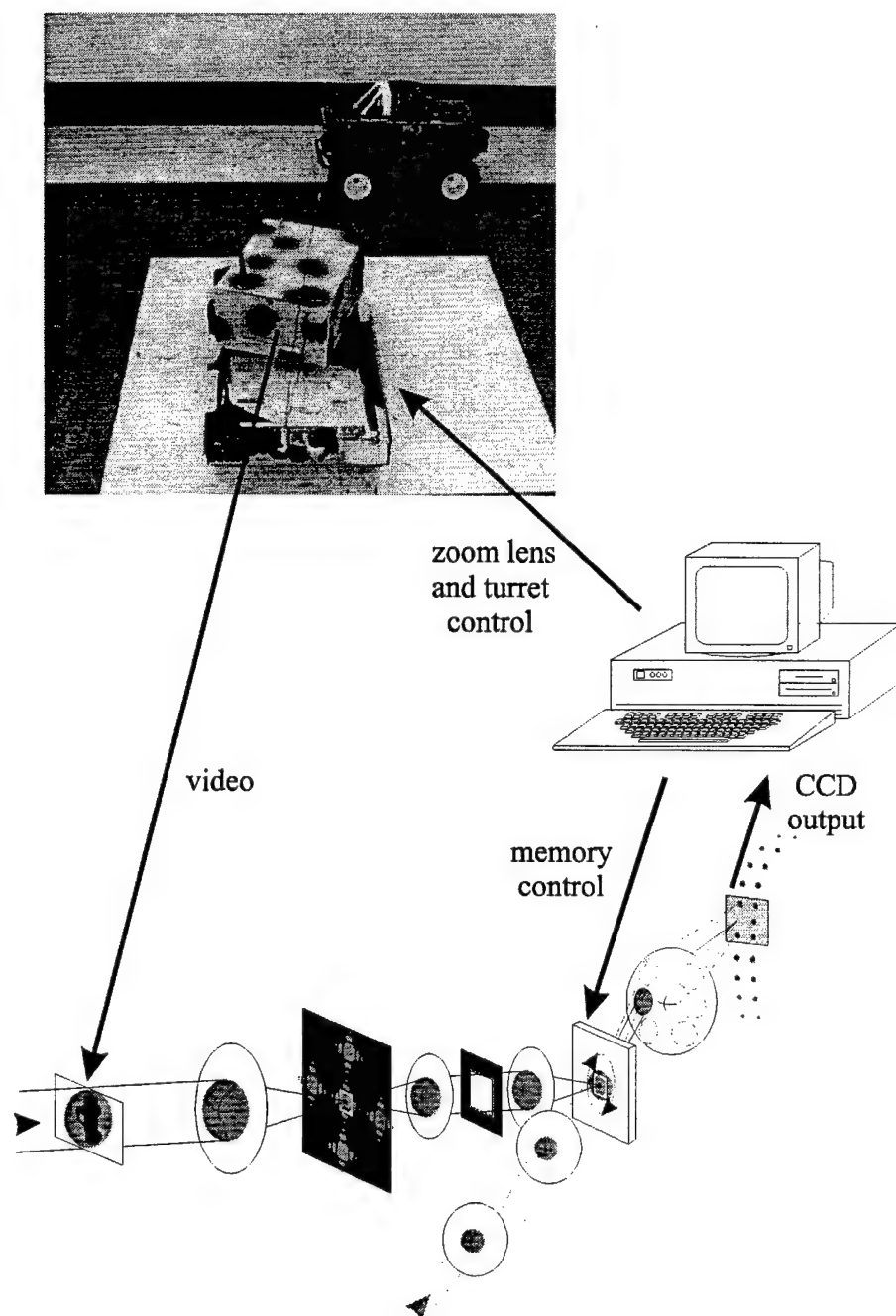
Figure 4.20 : An overview of the entire monster truck tracking system.

Since the 19 views of the target are continuous, they were stored sequentially using peristrophic multiplexing. The far and near views were stored in two separate angular positions as shown in Figure 4.18. The exposure intensity and recording time were similar to the autonomous vehicle navigation experiment discussed in Chapter 4.2.
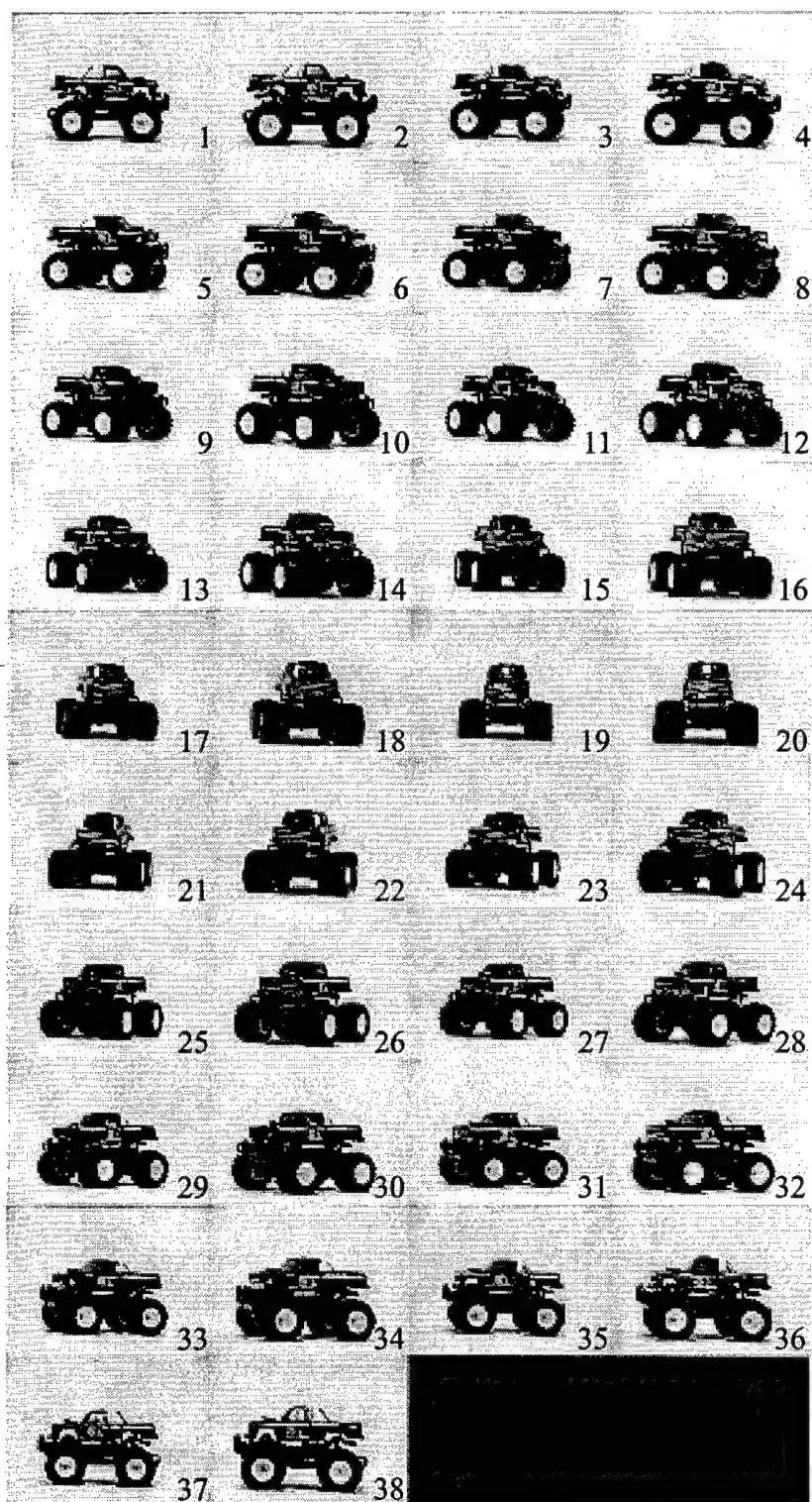
Figure 4.21 : 38 images of the red monster truck stored as filters in the holographic database.

4.3.2.3 Experimental Results

220

The experiment starts by placing the monster truck in the FOV of the tank, at one of the perspective stored in the holographic memory and within the scale defined by the near and far images. During the experiment, the tank was able to consistently keep the monster truck at the center of its FOV and at a constant scale regardless of movement, orientation, and scale changes. Figure 4.22 shows a sequence of frames of the tank tracking the monster truck. In each frame, the tank with its rectangular turret is shown in the foreground and the target in the background. In the lower left corner is the correlation plane (x' is vertical and y' is horizontal) and the lower right corner displays the video from the camcorder inside the turret (the tank's FOV). Frame 1 of Figure 4.22 shows the tank's turret turned to the port side to keep the target at the center of its FOV. As the target drives forward, the tank maintains a lock on the target by rotating the turret to the starboard side (frame 2). Frames 3 and 4 show the target moving closer to the tank. However, notice that the tank keeps the target at the same size in its FOV by adjusting the zoom lens of the camcorder inside the turret.

Since many perspectives of the target were stored in the holographic database, the tank can track the monster truck as it rotates. Frames 5 and 6 show the monster truck rotated by $10°$ and $20°$, respectively. The tank maintains a lock on the target by advancing or rewinding the holographic memory to find the best match. Frame 7 shows the target mostly obscured and no visible correlation peaks are present. In this case, the tank goes into a search mode and scans the horizon by rotating its turret. If the target comes back into its FOV, the tank locks on to the target and tracks it. Frame 8 shows that minor obscuration of the monster truck still produces bright enough peaks for the tank to track the target. The entire experiment was done in real-time and the maximum velocity of the

target in which the system was able to track was limited by the speed in which the turret

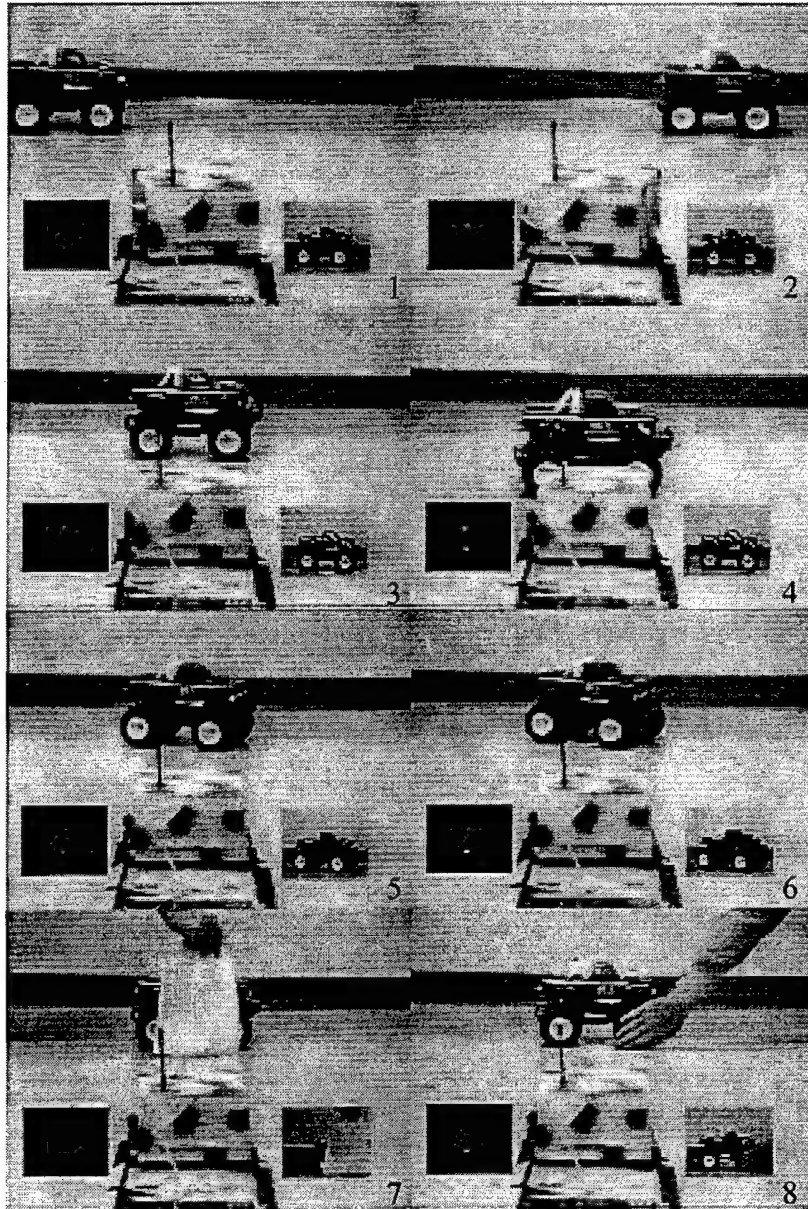can be rotated and how fast the holographic memory can be advanced or rewound.

Figure 4.22 : Sequence showing the tank tracking the monster truck.

4.3.3 Other Automatic Target Recognition and Tracking Experiments

In both the autonomous vehicle navigation and monster truck tracking experiments discussed previously, we exploited the continuity of the holographic database (waypoints or perspectives of the monster truck) to minimize the number of correlation peaks that had to be monitored at any given time and to simplify the control algorithm. However, it might not be possible to use this trick for all the scenarios in which an optical correlator might be deployed. We can generalize the optical correlator systems shown in Figure 4.4 and 4.17 by monitoring all the correlation peaks simultaneously as shown in Figure 4.23. Each hologram in the setup shown in Figure 4.23 is recorded with a distinct reference beam delivered by the reference beam deflection system. The reference beam is deflected along both x' and y' to generate a 2-D grid of correlation peaks at the correlation plane. The correlation peaks at the correlation plane have to be spatially separated enough so they do not enter each other's 'domain' as the input image is shifted. The size of this domain is defined by the shift invariance of the setup and the focal length of the Fourier transform lens right before the correlation plane. A CCD array can be placed at the exact correlation plane to monitor all the correlation peaks. However, since the aperture of the CCD array is usually very limited and we are interested in shift invariant correlators, only a few distinct peaks would fit on the CCD array. A better way is to place a white screen at the correlation plane and use a CCD camera to look at the entire white screen from a distance. This way, all the peaks on the white screen can be monitored simultaneously. The down side of doing it this way is now the peaks have to be more intense to be picked up by the CCD camera and the resolution of the peaks is reduced. In most cases these drawbacks are not important because the

photopolymer has very high dynamic range and we are only interested in relatively large shifts in the peaks' position.
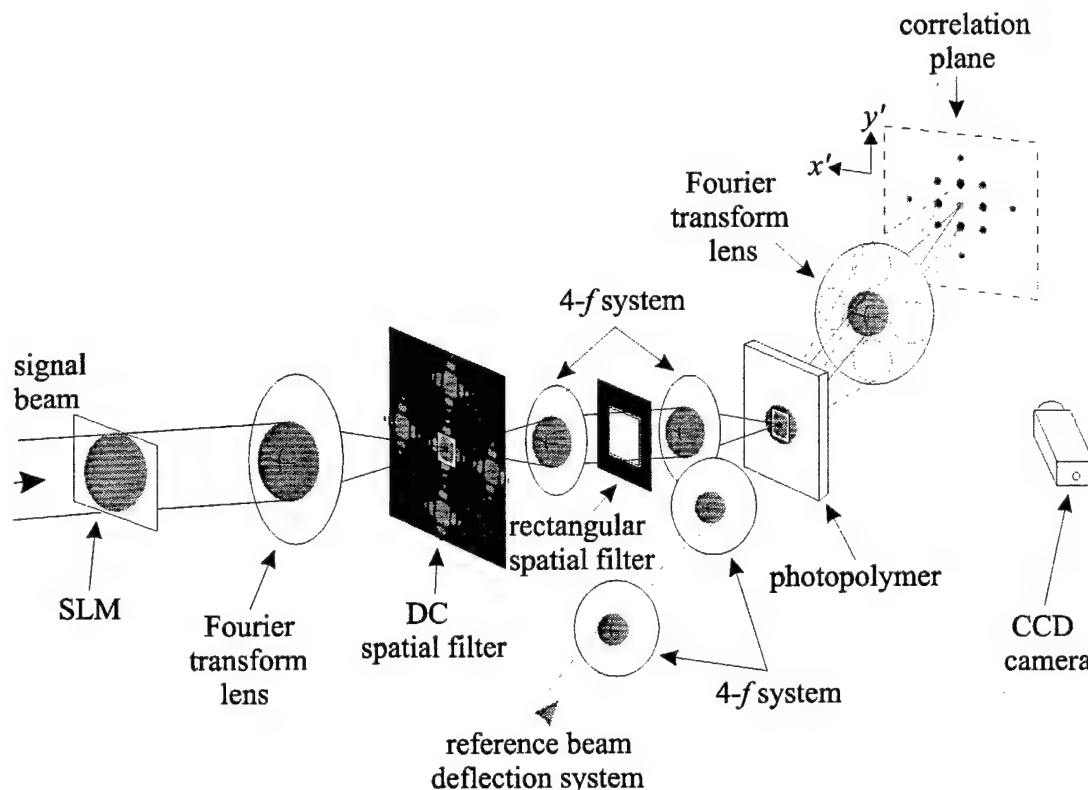

Figure 4.23 : The schematic diagram of a generalized optical correlator setup using holographic database.

The results of the cross-correlation between input images and the holographic database is displayed on the white screen in real-time. The CCD camera reads the white screen and a computer analyzes the video to find the intensity and position of all the correlation peaks. Depending on the organization of the holographic database and the goal of the system, a control algorithm can be devised to produce the correct response. One thing the control algorithm doesn't have to worry about in this case is advancing or rewinding the holographic database since the entire database is accessed at the same time.

4.3.3.1 Scud Missile Launcher

In both the autonomous vehicle navigation and monster truck tracking experiments, the system had to start from some initial condition. For the autonomous vehicle navigation experiment, the vehicle was limited to starting with the same view as frame 1 of Figure 4.11 or frame 1 of Figure 4.12. Similarly for the monster truck tracking experiment, the target had to be placed initially in the same perspective as shown in frames 1 and 2 of Figure 4.21 and at about the same scale. This is because only a subsection of the correlation peaks were monitored. Therefore, the input scene has to match with one of the filters in the holographic memory to generate a bright peak on the CCD array at the correlation plane before starting the system. However, if all the correlation peaks were monitored, then the computer can search for the brightest peak when the system is activated. This is nice because we cannot always have the target at some starting position for the optical correlator to lock onto.

For example, during the Gulf War, a major nuisance to the Allied Forces were the Scud missile launchers. Even though they were ineffective when conventionally armed (poor accuracy), they could reach Israel and cause major casualties if nuclear or biologically armed. The Allied Forces dispatched numerous planes to scour the desert, looking for Scud missile launchers but with little success. Perhaps an optical correlator system can prove useful in locating such targets and free the pilot to worry about other things.

Since it would be impossible to ask the Iraqis to always park the Scud missile launchers in a certain direction, it is necessary to store all the different perspectives of the launcher in the holographic database. Furthermore, the perspective of the launcher when

it comes into the FOV of the camera mounted on the plane is unpredictable. Hence we would need a system where all the correlation peaks can be monitored simultaneously.

Figure 4.24 shows 72 different perspectives of a model Scud missile launcher we stored as filters in an optical correlator system similar to the one shown in Figure 4.23. The control algorithm to determine if a Scud missile launcher is in the FOV of the camera is very simple. The computer locates the brightest peak and if the intensity of the peak is greater than some preset threshold, then the computer reports that a Scud missile launcher is found. The computer also reports the heading (perspective) of the launcher based on the position of the brightest peak in the correlation plane. Each peak in the correlation plane corresponds to a filter in the database. From the position of the brightest peak, the computer knows which filter has the best response to the input image and therefore the heading of the target.

The frames in Figure 4.25 show the system in action. Frames 1 through 4 show that the system can locate the launcher from many different perspectives. The upper section in each of the four frames reports "SCUD LAUNCHER FOUND : BEARING" 10, 120, 200, and 310 degrees respectively. Frame 5 shows that significant portions of the launcher can be obscured by trees and the system can still locate the launcher. However if the trees block too much of the launcher, then the system would lose the target and report "SEARCH...SEARCH...SEARCH...SEARCH..." as shown in Frame 6. Frame 7 shows that if the trees are in the background, then the system can still locate the launcher since the edges of the launcher remain sharp. Blocking all the wheels of the launcher with smaller trees did not make the system lose track of the target as shown in frame 8.

'stuff' that is in the FOV, the longer the bar is from left to right. The area below the complexity bar is the live video from the camera.
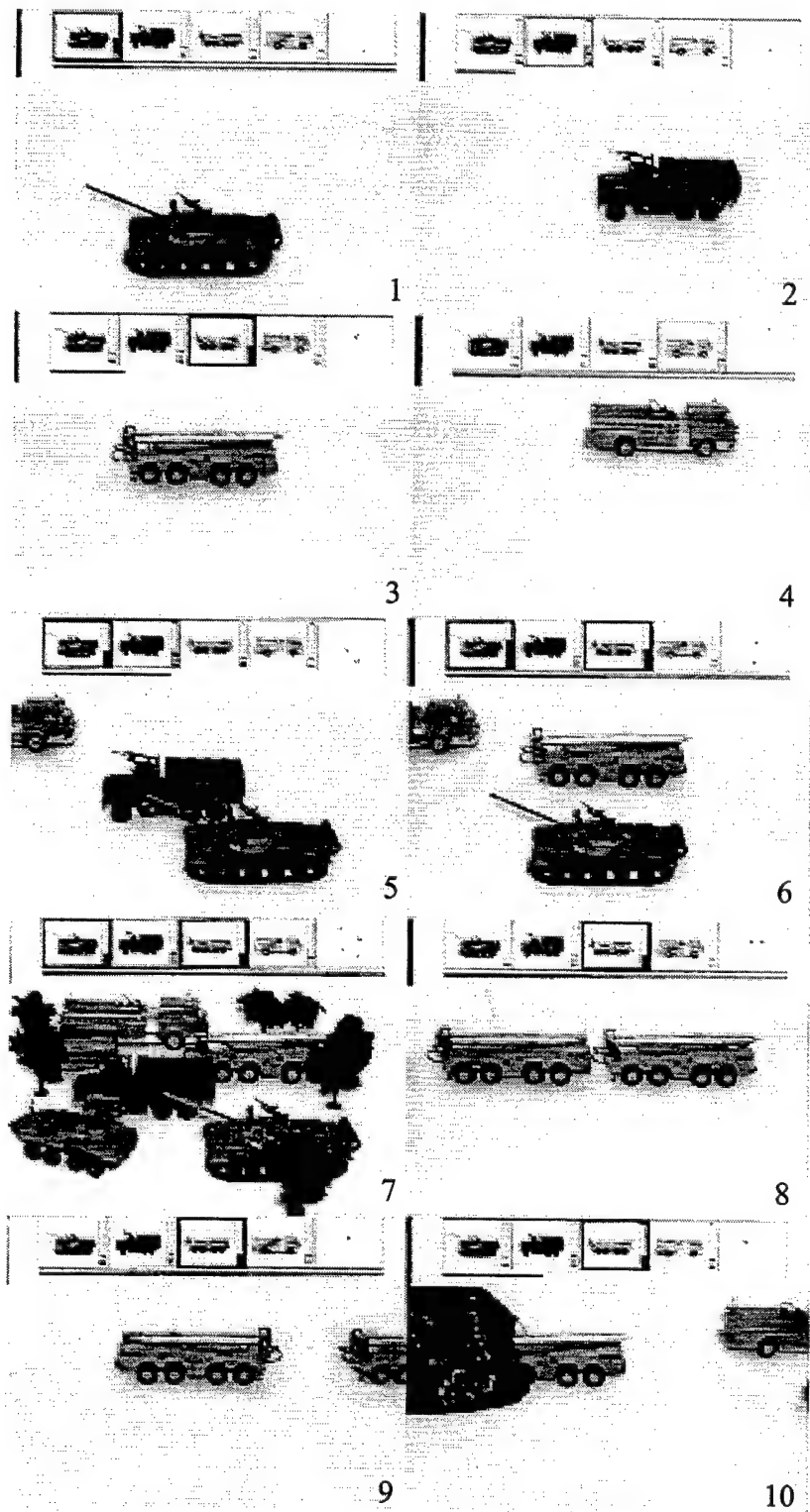
plane by deflecting the reference beam in both x' and y' during storage. If the intensity of anyone of these peaks passes threshold during cross-correlation, then the corresponding target is in the FOV of the camera. Furthermore, from the position of the peak, the target can be tracked as it moves across the FOV of the camera.
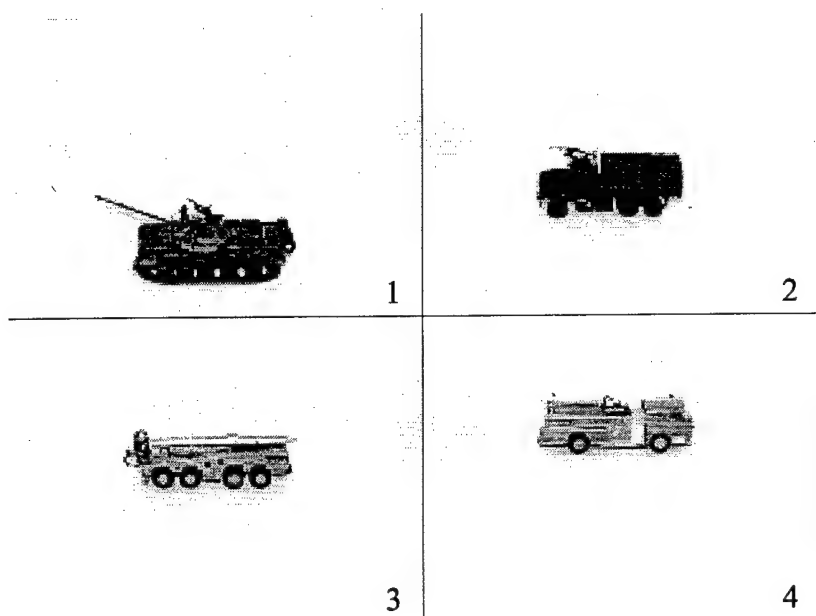

Figure 4.26 : Four different targets we would like to track simultaneously with an optical correlator system.

The frames in Figure 4.27 show the system in action. In each frame, the upper section have four miniature pictures of the stored targets, each with a cross-correlation intensity bar next to the picture. The miniature pictures are arranged from left to right, in the order of tank, troop transport, Scud missile launcher, and fire truck. When the peak intensity bar passes some preset threshold, that means the target is in the FOV of the camera. This is indicated by outlining the picture of the corresponding target with a box and displaying a small dot in the upper right corner to show the position of the target in the FOV. There is another energy bar below the four miniature pictures in the upper section. This bar shows the complexity of the scene in the FOV of the camera. The more
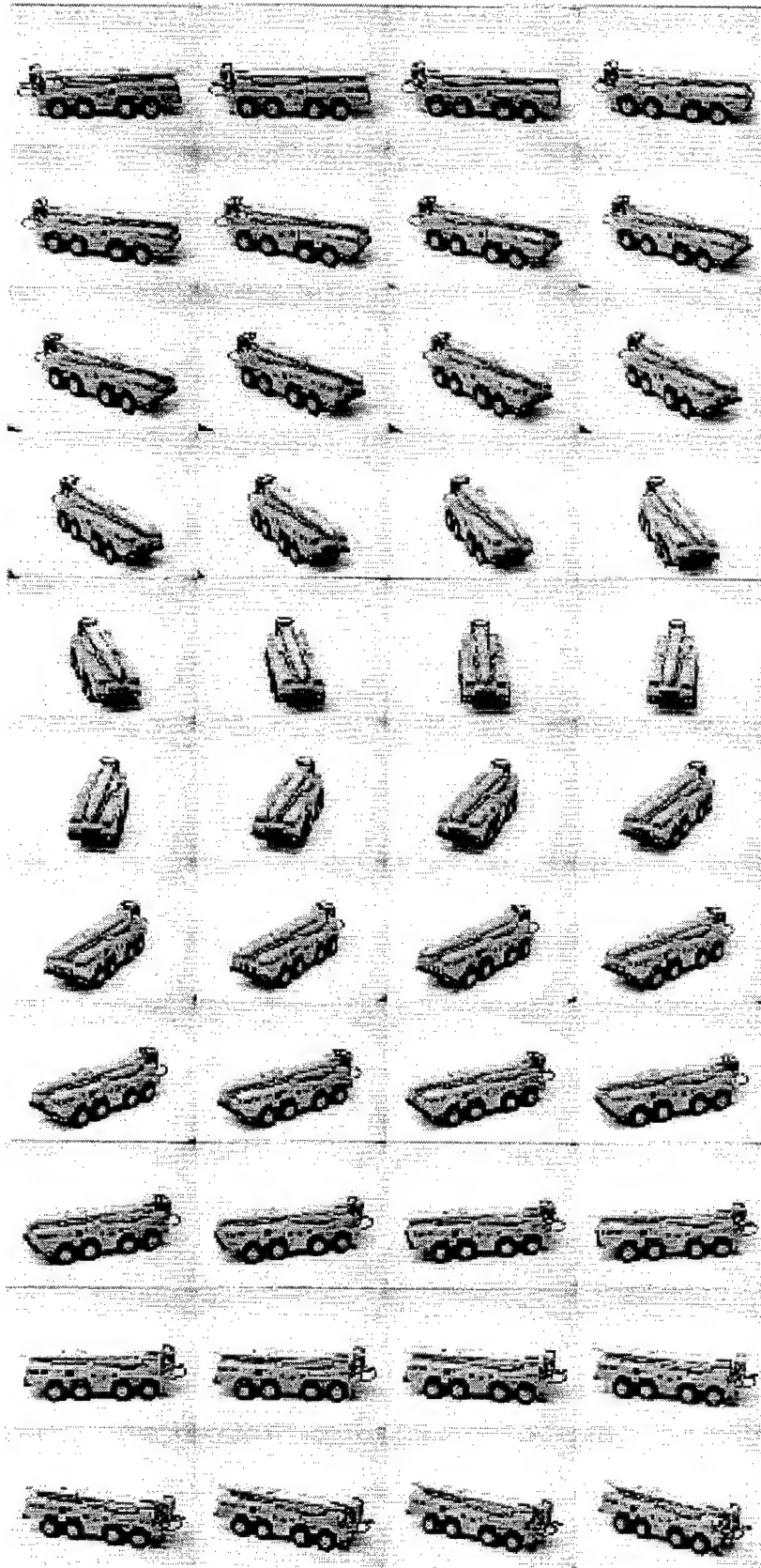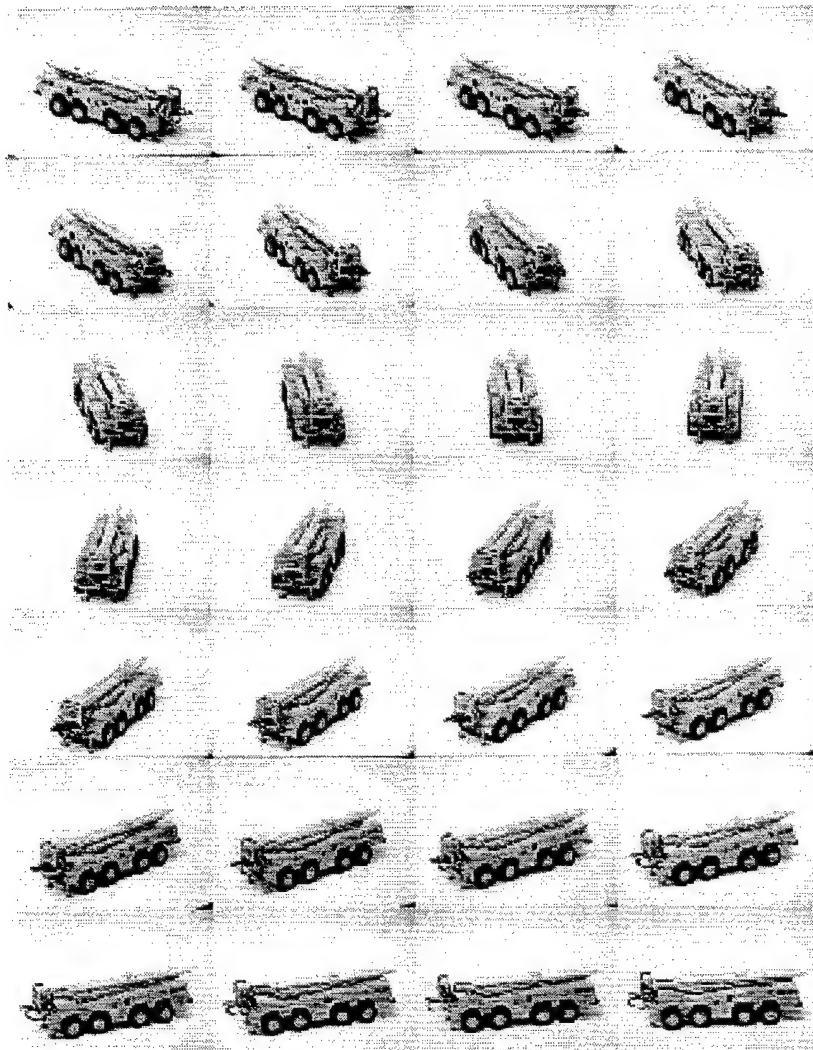
Figure 4.24 : 72 images of Scud missile launcher stored as filters in the holographic database.

Frame 9 of Figure 4.25 shows that the system can still find the launcher if the missile was removed. For that particular perspective of the launcher, the missile accounted for less than 20% of the edges and therefore did not reduce the peak intensity too much when removed. However, for some other perspectives, the missile might account for a large percentage of the total edges and the system would lose the launcher as shown in frame 10. Frame 11 shows the missile in the launch position. The system

can location the launcher from perspectives where the missile and the launch stand account for a small percentage of the total edges.
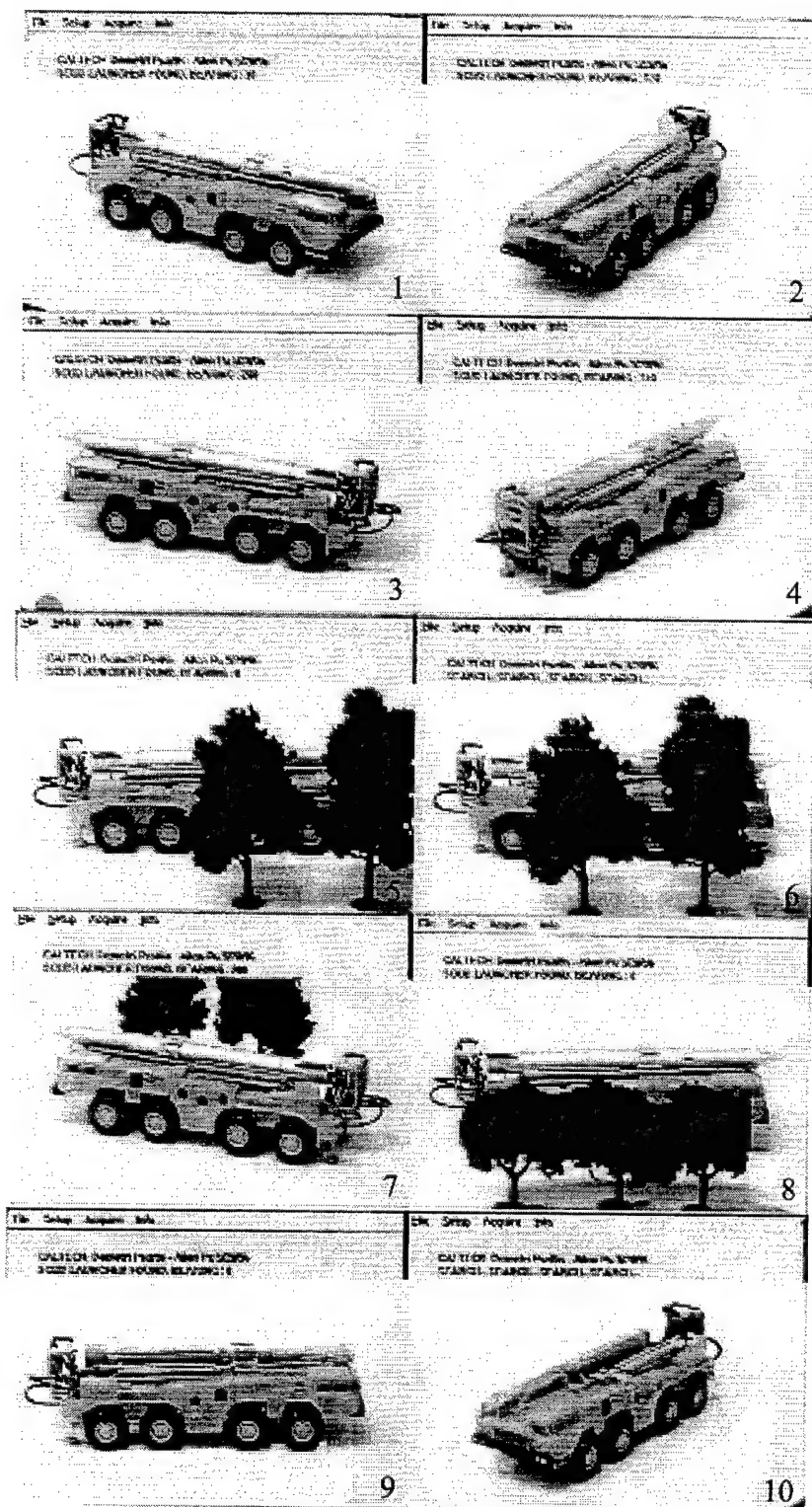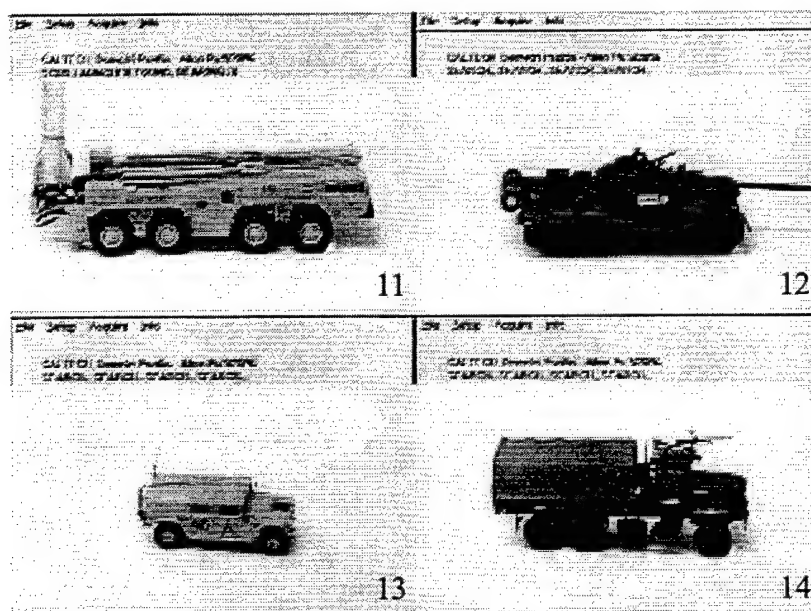
Figure 4.25 : Sequence showing the Scud missile launcher experiment.

Frames 12, 13, and 14 of Figure 4.25 shows that the system is not fooled by targets that do not look like the Scud missile launcher. This is because none of the peaks passed threshold when a T-72 tank, HUMMER, and troop transport were placed in the FOV of the camera.

4.3.3.2 Multiple Targets

Another scenario in which the filters in the holographic database have no continuity is when multiple targets have to be tracked simultaneously. For example, Figure 4.26 shows four different targets that we would like to track with an optical correlator system. One or more of the four targets can appear in the FOV of the camera at any given time so the system has to monitor all four peaks simultaneously. We used the generalized optical correlator setup shown in Figure 4.23 for this experiment. The four targets shown in Figure 4.26 were organized as a 2 × 2 grid of peaks at the correlation
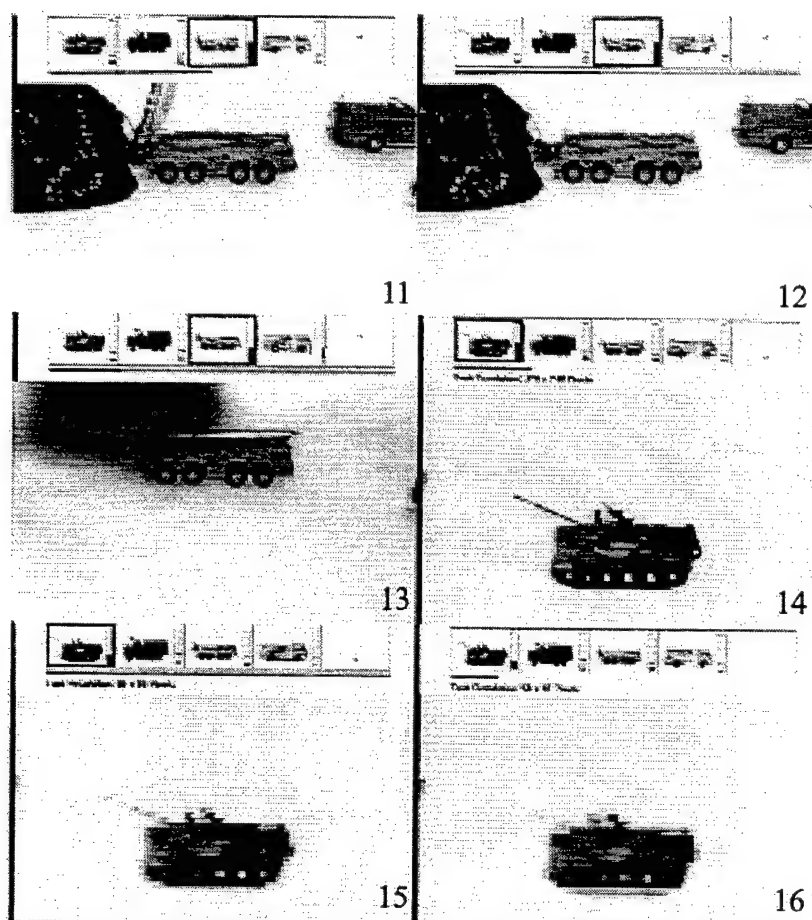
Figure 4.27 : Sequence showing the optical correlator system tracking multiple targets.

Frames 1 through 4 of Figure 4.27 show that each target can be recognized and tracked when put in the FOV of the camera. In addition, multiple targets can be tracked simultaneously as shown in frames 5 and 6. Frame 7 shows that all the targets can be recognized even in the presence of clutter (trees, other vehicles) and some minor obscuration. Due to the shift invariance of the correlation function, the system can pick out multiple targets of the same type. Frame 8 shows that the system recognizes two Scud missile launchers in a convoy (see the two dots in the upper right hand corner). Since the Scud missile launcher has good symmetry, the system can still recognize the launcher when it is rotated by 180° as shown in frame 9. Frames 10, 11, and 12 show the

Scud missile launcher coming out from a bunker and firing the missile. Once approximately 60% of the Scud missile launcher is in plain view, the system recognizes the launcher. The system has enough edges from the body of the launcher to track the launcher even when the missile has been fired and the launch stand in the down position. Frame 13 shows that the launcher can still be recognized in poor lighting conditions such as shadows from clouds, as long as most of the edges are still visible. The targets were stored in the holographic database at full resolution of 640×480 pixels per image. Frame 14 shows the tank at $1/4^{th}$ of the original resolution (320×240) and the system has no trouble recognizing the target. Even down to $1/64^{th}$ of the original resolution (80×60), the tank can still be recognized (frame 15). The correlation peak intensity for the tank drops below threshold for resolutions at or less than 58 by 43 pixels (frame 16)

4.3.3.3 Daryl the Crash Test Dummy

In all the automatic target recognition and tracking experiments we have shown in this chapter, the filters in the holographic database are just images of the target taken with a video camera. To track a target from many perspectives and scale, a lot of filters would have to be stored (number of filters = number of perspectives times the number of scales). We can reduce the number of filters required by first normalizing the scale of the input object to a fixed size. Then in the holographic database, only the different perspectives of the target at one scale is needed.

For example, we setup an experiment to track human heads. A video camera with zoom control was mounted on a pan-tilt unit and placed in the hallway to track people walking by. The objective of the system is to normalize the heads of the people in the

FOV of the video camera and see if one of them is Daryl (one of the Crash Test Dummies). Figure 4.28 shows the filters stored in the holographic database of an optical correlator system for this experiment. Frames 1 through 8 of Figure 4.28 are used to track human heads and to normalize it to one scale by changing the zoom of the video camera. These filters were generated with a computer and frame 5 is the desire scale. Frames 9 through 17 are used to see if the input head is Daryl.
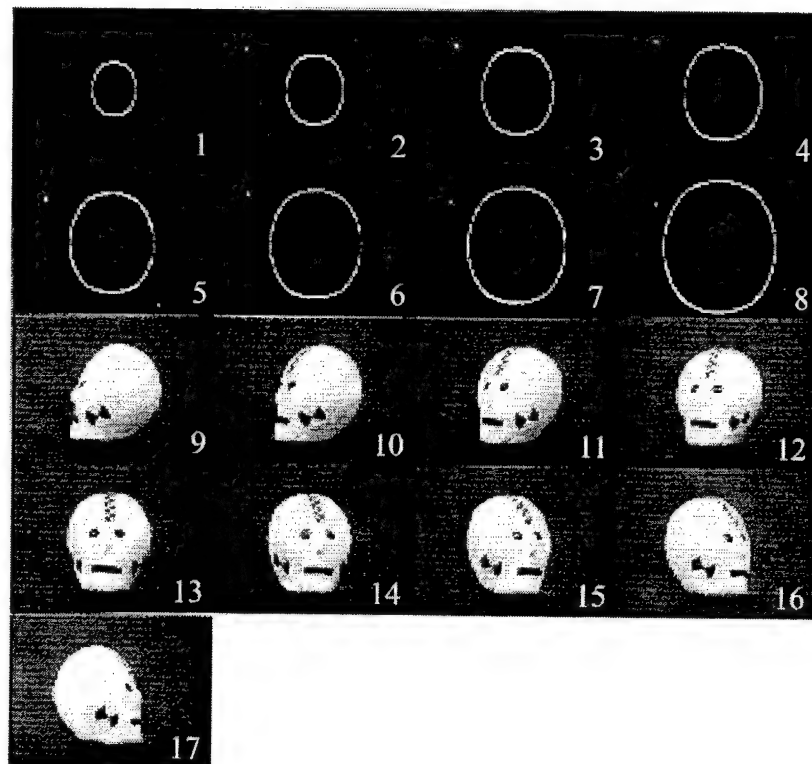

Figure 4.28 : Images stored in the holographic database to recognize and track Daryl.

The sequence of frames in Figure 4.29 shows the system in action. Frames 1 through 4 show that the input head has been normalized and it is Daryl. The upper left hand corner in each frame reports the direction Daryl's head is facing in degrees and the direction Daryl is moving with little arrows. The system keeps Daryl's head at the center

of its FOV by rotating the video camera with the pan-tilt unit and at the desired scale by
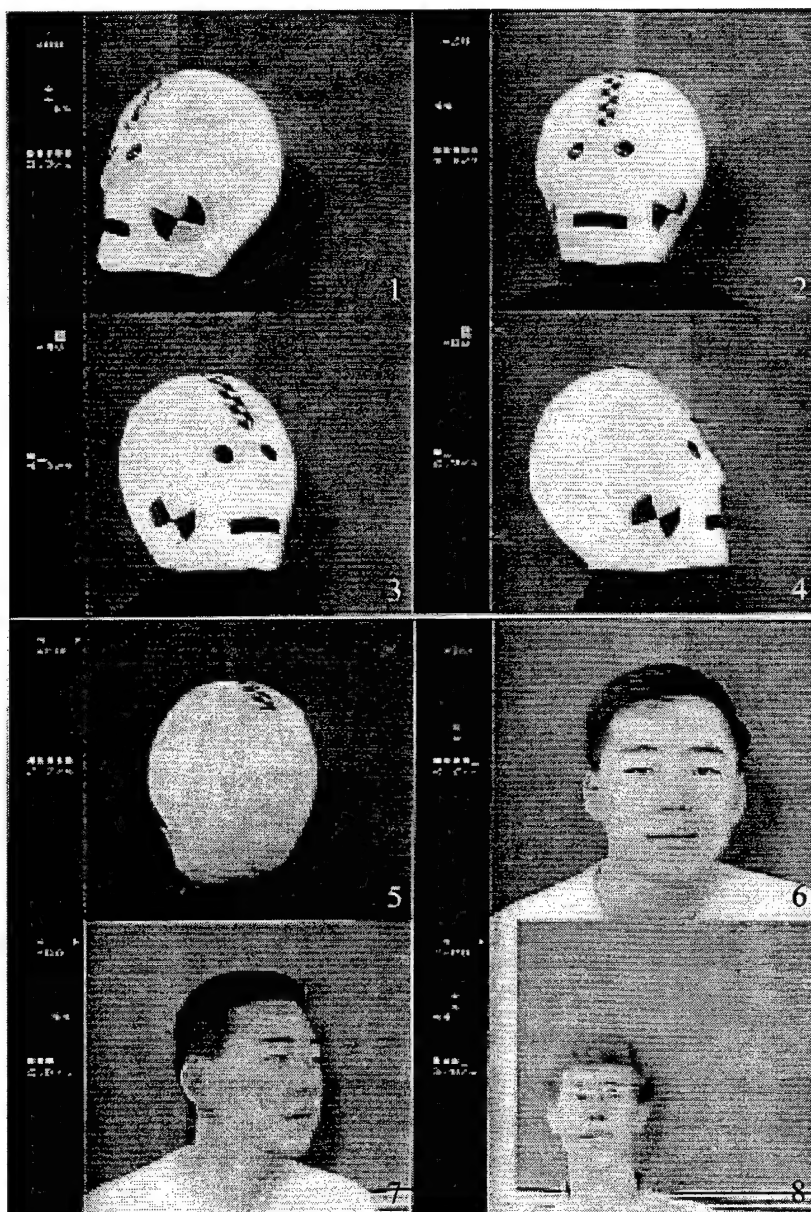
controlling the zoom.

Figure 4.29 : Sequence showing the optical correlator system tracking heads.

Frames 5, 6, and 7 shows that the system normalizes and tracks human heads even

if it doesn't look like one of the stored perspectives of Daryl. After the system is sure that

the head in its FOV is not Daryl, the system would stop tracking and wait for another

head to enter its FOV. However, this system was not able to normalize and track all human heads. Frame 8 shows that the system failed to normalize Greg Billock's head. Notice his head is sort of flat on top and therefore the elliptical filters couldn't locate his head in the input scene. Perhaps the elliptical filter can be improve to normalize and track Greg's head as well.

For this experiment, instead of having to store 72 different filters to track Daryl (9 perspectives at 8 different scales), only 17 filters were required (8 elliptical filters to normalize the input head and 9 matched filters to determine if the head is Daryl's).

4.4 User Identification Through Sequential Fingerprint Inputs

4.4.1 Introduction

Currently there are many fingerprint verification and identification systems that are being developed or are already on the market. A large majority of these systems are for verification only. The user enters a Personal Identification Number (PIN) or inserts an ID card before entering his or her fingerprint. The PIN allows the system to retrieve information about the user's fingerprint from a database for verification against the entered fingerprint. In the case of an ID card, it is possible to encode the user's fingerprint information on the card's magnetic strip or a microchip. The system then simply verifies the fingerprint retrieved from the ID card against the entered fingerprint.

A much harder problem to solve is fingerprint identification where the user is identified solely through the entered fingerprint. The system has to figure out who the user is by comparing the entered fingerprint against a database. If the database is large, then the search time can be quite long. Furthermore, it is very likely that several

fingerprints in the database match well enough with the entered fingerprint to be in doubt. This is especially true for fingerprints since moisture, oil, pressure applied, and movement can alter the registration, shape, size, and structure of an input fingerprint. To achieve a good identification rate (near 100%), often a complex and computationally intensive algorithm has to be applied with diminishing returns. Therefore, very few systems are capable of identifying a user positively through a single fingerprint. In this section, we describe an optical correlator system that can identify users through sequential fingerprint inputs.

4.4.2 Theory of Operation

For user identification, an optical correlator can cross-correlate an input fingerprint with a database of stored fingerprints very quickly, so the search time is not a problem. From the strength of the cross-correlations, the user can be identified. However, in order to prevent false identification (intruders recognized as in the database or one user recognized as another), each filter (fingerprint) in the database must have a threshold value for which the cross-correlation intensity must be greater than to trigger a response. If the threshold is set too high, then authorized users might have a hard time getting the system to identify them. On the other hand, if the threshold is set too low, then false identification might occur.

The primary goal of a security device is to restrict access to those authorized. However, if that means an authorized user has to waste a lot time getting the system to identify him or her, then the system is not going to be very useful. Luckily since optical correlators can compute in real-time, even when the thresholds are set high, the user can

continuously move his or her finger until the corresponding peak intensity passes threshold (movement can sometimes correct rotation and position errors). Nevertheless from the users' point of view, it is much better to have low thresholds so they can get past the system quickly. One way to lower the thresholds and still have a low false identification rate is to use multiple fingerprints from each user. Each user has to enter several fingerprints in order to be identified by the system. The probability of different users having several 'similar' fingerprints is fairly small. On the other hand, this method increases both the memory requirement and computational time for user identification. Since optical correlators can compute in real-time and the holographic database can be very large, the impact of using several fingerprints from each user on the system's performance is minimal.

If several fingerprints have to be entered to gain positive identification, it is possible to play around with the sequence in which the fingerprints are entered. To increase security and to minimize false identification, we can have the users enter his or her fingerprints in a secret sequence, like a PIN. For example, a system can be implemented to recognize three different fingerprints from each user (i.e. the middle three fingers on the left hand) and to receive a five fingerprint long sequence (constructed from the three fingerprints). Figure 4.30 shows the flowchart of the user identification algorithm using sequential fingerprint inputs. To gain recognition from the system, a user enters his or her fingerprints sequentially as if he or she is keying a PIN at the cash machine (the fingerprints are read into the system through a fingerprint input device). The system first correlates the entered fingerprint against the database to see if it belongs to one of the authorized users. If it is in the database, the system then identifies which

user the fingerprint belongs to and registers the fingerprint along with the sequence number in the user's temporary stack memory (i.e. middle finger recognized as belonging to Allen Pu and is registered in Allen Pu's stack memory as input number 3). When the threshold is set low (for quick fingerprint recognition), it might be possible for the intensity of several correlation peaks to be greater than the threshold. If the peaks correspond to different fingerprints of the same user, then the brightest one wins. If the peaks correspond to different users, then the fingerprint along with the sequence number is registered in each user's stack. Once the system has recognized 5 different fingerprints (not necessarily all belonging to the same user), the system compares each authorized user's stack against his or her respective secret PIN to see if there is a match. Under optimal conditions, all the stacks should be empty except one and the fingerprint sequence in the stack matches the secret PIN of that user. In that case, the user is positively identified. On the other hand, if none of the stacks are completely filled after 5 recognized fingerprint inputs or if the sequence in the stack does not match with that user's PIN, then the user is not positively identified and rejected.

The above algorithm minimizes false identification three ways: (1) fingerprint from the user is compared with a database for identification; (2) three different fingerprints from the user is required to generate the five fingerprint sequence (redundancy); (3) the fingerprint sequence keyed has to match the secret PIN known only by the authorized user. Not only does this system minimize false identification, it is also tamper-resistant. In order to be identified by the system, an intruder must provide the three different fingerprints of a user that is in the database. Furthermore, the intruder must enter the fingerprints in the order of the secret sequence of that user. For example,

even if the PIN of an authorized user is compromised (through peeping), the intruder would still have no way of entering the sequence. Vise versa, if the fingerprints of a user is compromised, the intruder would still have to obtain the PIN of that user.
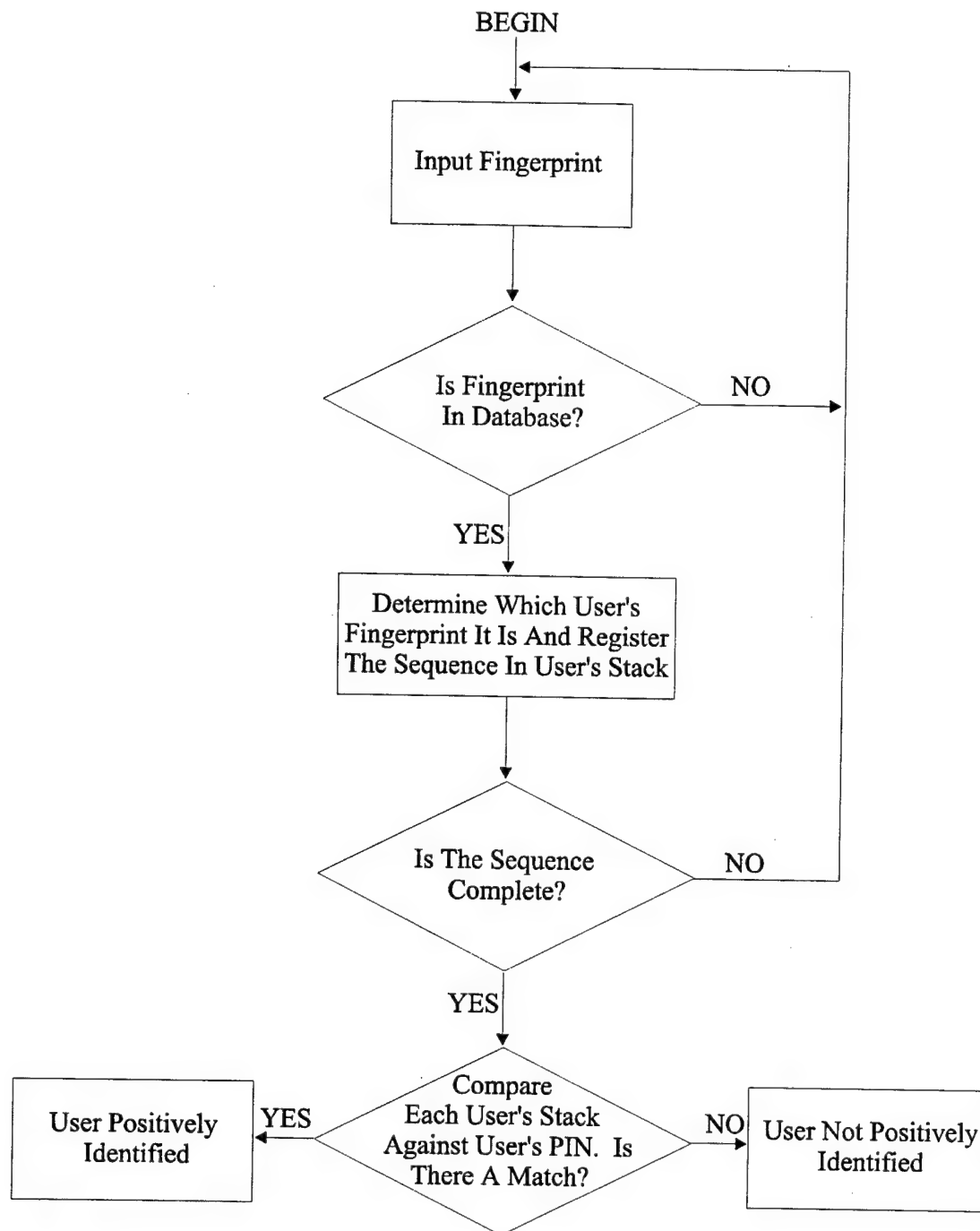

Figure 4.30 : The flowchart of the user identification algorithm.

4.4.3 Experimental Setup and Holographic Database

The optical correlator setup used in this experiment is the same as the one shown in Figure 4.23. Three fingerprints from each of the five 'volunteer' graduate students were stored in the holographic database. At the correlation plane, the correlation peaks were organized as a 3 × 5 grid by deflecting the reference beam in both x' and y' during storage. Each row represented a different user and each column represented a different fingerprint. For security reasons, the fingerprints stored will not be reprinted here.

Figure 4.31 shows the overview of the entire user identification system. An user enters his or her fingerprint into the optical correlator through an Identix fingerprint reader. The Identix fingerprint reader outputs RS-170 video which can be directly connected to the TVT-6000 SLM. The result of the cross-correlation between the input fingerprint and the database is read into a computer by a CCD camera and a framegrabber board. Depending on the intensity of the peaks and the preset threshold, a fingerprint is either recognized (one of the five yellow LEDs is turned on at the display console) or rejected (nothing happens). If the entered fingerprint is recognized, then the computer figures out which user's fingerprint is entered and updates the corresponding stack memory to form the fingerprint sequence. Once five fingerprints have been recognized, the computer makes the final decision by comparing the entered PIN with the secret PIN and either displays the user's name (turning on one of the five green LEDs at the display console) or rejects the user (a red LED is made to blink for several seconds). Notice the computer does not have to advance or rewind the holographic memory since all the correlation peaks are monitored simultaneously.
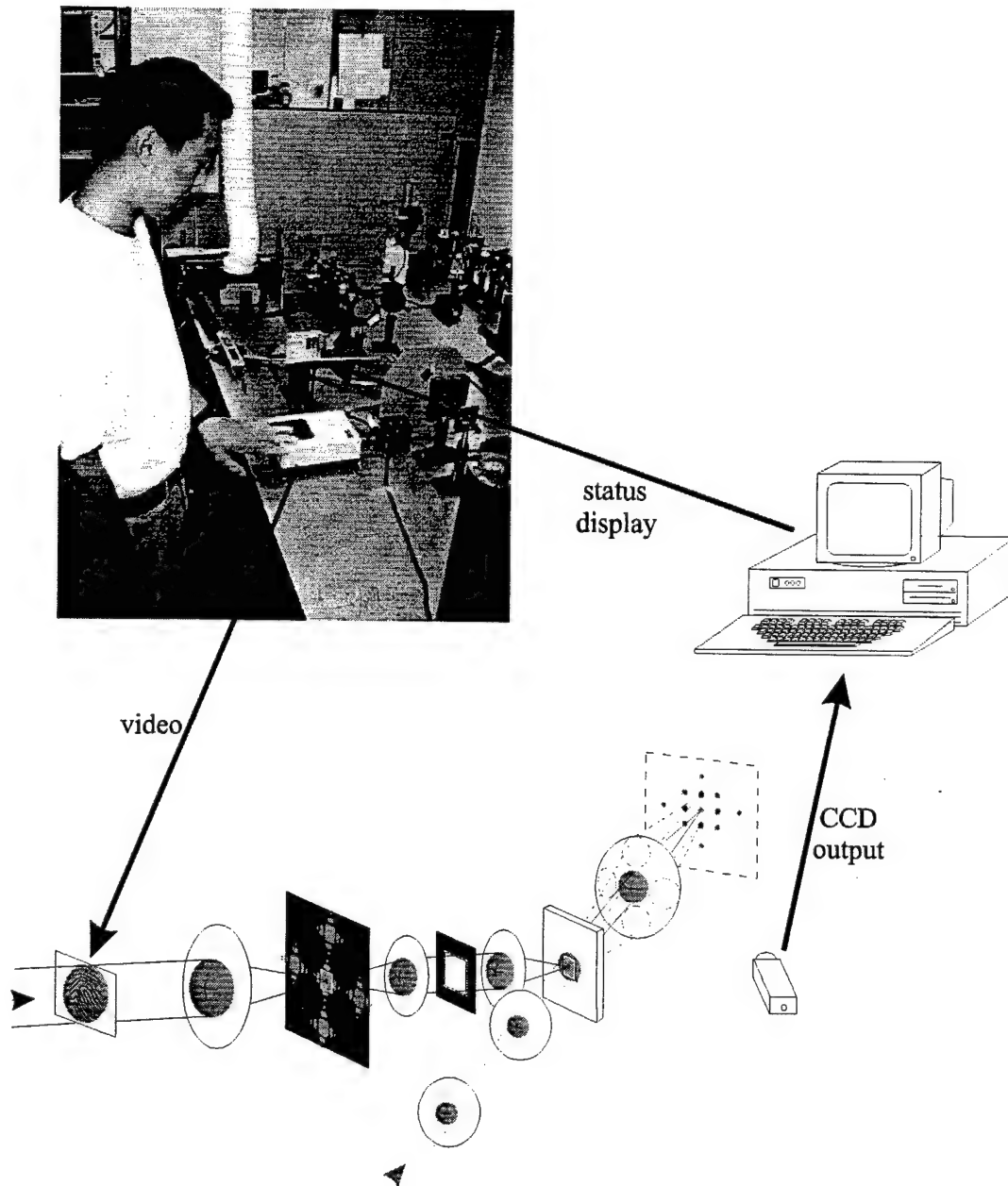
Figure 4.31 : An overview of the entire user identification system.

4.4.4 Experimental Results

Once the fingerprints have been stored and PINs selected, different users in the database are always correctly identified when the proper sequence for each user is entered. Furthermore, intruders whose fingerprints are not in the data base are always

rejected. The entire identification process takes less than 5 seconds with most of the time spent on switching fingers for input. Figure 4.32 shows a sequence from the user identification experiment. In each frame, the insert in the lower left corner shows a blow up of the Identix fingerprint input device and the status display panel. Frames 1 through 5 show one authorized user entering his sequence of fingerprints to gain access. In each frame, one extra LED lights up on the status display panel indicating the acceptance of the entered fingerprint. In frame 5, the sequence is complete and a LED lights up in the second row of the status display panel indicating that the PIN matches the entered fingerprints and the user is identified as Allen Pu. In frame 6, another authorized user enters his sequence of fingerprints and a different LED in the second row of the status display panel lights up, identifying the user as Robert Denkewalter. Frame 7 shows an unauthorized user trying to enter his fingerprints into the system. However, none of his fingerprints were able to trigger a response from the system. Frames 8 and 9 show another unauthorized user entering fingerprints into the system. This time, he is able to trigger a response by using his pinky finger. After entering 5 fingerprints, the system rejects him because the sequence does not match the fingerprints entered. It is possible that the unauthorized user had triggered the same fingerprint in the database all 5 times, generating a sequence of 11111, 22222, or 33333. These types of PINs are automatically rejected since they do not consist of three different fingerprints. It is also possible that the unauthorized user had triggered fingerprints from several different authorized users. In this case, none of the authorized users' stack is full and the incomplete PIN is rejected.

Figure 4.32 : Sequence showing the user identification experiment.

Frame 10 of Figure 4.32 shows a desperate unauthorized user kidnapping the hand of an authorized user in an attempt to gain access. However, as long as the authorized user does not disclose his PIN, the unauthorized user cannot gain recognition from the system and the system shuts down after a few attempts.

4.5 Discussions and Conclusions

The main goal of these experiments is to demonstrate that a simple optical correlator system using a holographic database can solve problems that are complex and/or time consuming to solve with conventional electronic computers. The optical correlator systems shown in this chapter were able to in real-time - guide a vehicle through a pre-programmed path, automatically recognize and track targets, and identify users through fingerprints. In each experiment, less than 100 holograms were required, which is only a small fraction of the holographic database's total capacity. Up to 1,000 holograms have been stored at a single location using the DuPont 100 micron thick photopolymer (Chapter 2.4.5) and up to 10,000 holograms have been stored in photorefractive crystals [19]. Even larger capacity can be achieved by multiplexing holograms spatially on a disk [20]. Such a large capacity can allow an entire building or a campus to be mapped for autonomous vehicle navigation, a number of targets stored from all possible perspectives, scales, and elevations for automatic target recognition and tracking, the fingerprints from the employees of an entire company stored for security devices. Furthermore, the optical correlator and the holographic database can be miniaturized to fit in small places such as onboard a vehicle, missile or in the walls of a building as self contained units.

References

[1] A. Kozma and D. L. Kelly. "Spatial filtering for detection of signals submerged in noise," *Applied Optics*, 4:387, 1965.

[2] A. B. Vander Lugt. "Signal detection by complex spatial filtering," *Radar Lab., Rept. No.* 4594-222-T, Institute of Science and Technology, The University of Michigan, Ann Arbor, 1963.

[3] J. T. Tippett, et al. *Optical and electro-optical information processing*, pp:133-139, The M.I.T. Press, Cambridge, Mass., 1965.

[4] J. W. Goodman. *Introduction to Fourier Optics*, pp:171-179, McGraw-Hill, 1968.

[5] H.-Y. S. Li. *Photorefractive 3-D disks for optical data storage and artificial neural networks*. Ph.D. thesis, California Institute of Technology, 1994.

[6] C. S. Weaver and J. W. Goodman. "A technique for optically convolving two functions," *Applied Optics*, 5:1248-1249, 1966.

[7] X. A. Shen and R. Kachru. "High speed pattern recognition by using stimulated echoes," *Optics Letters*, 17(7):520-523, 1992.

[8] C. Gu, J. Hong, and S. Campbell. "2-D shift-invariant volume holographic correlator," *Optics Communications*, 88:309-311, 1992.

[9] D. Psaltis, M. A. Neifeld, and A. Yamamura. "Image correlators using optical memory disks," *Optics Letters*, 14(9):429-432, 1989.

[10] K. Curtis and D. Psaltis. "3-dimensional disk-based optical correlator," *Optical Engineering*, 33(12):4051-4054, 1994.

[11] F. T. S. Yu and D. A. Gregory. "Optical-recognition - architectures and techniques," *Proceedings of the IEEE*, 84(5):733-752, 1996.

[12] B. Javidi, J. Wang, and Q. Tang. "Nonlinear joint transform correlators," *Pattern Recognition*, 27(4):523-542, 1994.

[13] K. Sayano and F. Zhao. "Optical correlation of phase- and amplitude-encoded objects using wavelength multiplexed holographic storage elements," *International Symposium on Optical Memory and Optical Data Storage*, JTuD11, Maui, Hawaii, July 1996.

[14] E. D. Dickmanns and N. Muller. "Scene recognition and navigation capabilities for lane changes and turns in vision-based vehicle guidance," *Control Engineering Practice*, 4(5):589-599, 1996.

[15] N. P. Papanikolopoulos, P. K. Khosla, and T. Kanade. "Visual tracking of a moving target by a camera mounted on a robot - a combination of control and vision," *IEEE Transactions on Robotics and Automation*, 9(1):14-35, 1993.

[16] U. M. Fayyad, P. Smyth, M. C. Burl, and P. Perona. "Learning approach to object recognition: applications in science image analysis," *Early Visual Learning*, U. M. Fayyad, Editor, Oxford University Press, 1996.

[17] B. V. K. Vijaya Kumar. "Minimum-variance synthetic discriminant functions, " *JOSA-A*, 3(10):1579-1584, 1986.

[18] T. D. Hudson, A. S. Kransteuber, and D. K. McMillen. "Optical implementation of composite filters as volume holograms in $LiNbO_3$ for pattern recognition," *International Symposium on Optical Memory and Optical Data Storage*, JTuD5, Maui, Hawaii, July 1996.

[19] G. W. burr, X. An, F. H. Mok, and D. Psaltis. "Large-scale rapid access holographic memory," *Optical Data Storage Meeting 1995*, SPIE Vol:2514-39.

[20] K. Curtis and D. Psaltis. "3-dimensional disk-based optical correlator," *Optical Engineering*, 33(12):4051-4054, 1994.